

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 21:48:41 ; Search time 2370 Seconds
(without alignments)
8094.610 Million cell updates/sec

Title: SEQ1-96535C

Perfect score: 300

Sequence: 1 ggcattgacaggctaaatgc.....atctgctgattaggaagtat 300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

GenEmbl.*

1: gb_env.*

2: gb_pat.*

3: gb_ph.*

4: gb_pl.*

5: gb_px.*

6: gb_ro.*

7: gb_sts.*

8: gb_sy.*

9: gb_un.*

10: gb_vi.*

11: gb_ov.*

12: gb_hcg.*

13: gb_in.*

14: gb_om.*

15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 298.4 | 99.5 | 121896 | 5 | AL138713 |
| 2 | 143 | 47.7 | 281593 | 12 | AC172613 |
| 3 | 51.6 | 17.2 | 175961 | 6 | AC124705 |
| 4 | 51.6 | 17.2 | 184477 | 6 | AC117638 |
| 5 | 44.2 | 14.7 | 246893 | 12 | AC129383 |
| 6 | 44.2 | 14.7 | 314404 | 12 | AC130228 |
| 7 | 43.8 | 14.6 | 218961 | 12 | AC153067 |
| 8 | 43.6 | 14.5 | 4601 | 13 | DMU11584 |
| 9 | 43.6 | 14.5 | 19517 | 13 | DMU37541 |
| 10 | 43 | 14.3 | 173903 | 6 | AC163669 |
| 11 | 43 | 14.3 | 196140 | 12 | AC084107 |
| 12 | 42 | 14.0 | 154779 | 6 | AC137708 |
| 13 | 42 | 14.0 | 212071 | 6 | AC136754 |
| 14 | 42 | 14.0 | 223110 | 6 | AC123752 |
| 15 | 42 | 14.0 | 224009 | 12 | AC131058 |
| 16 | 41.8 | 13.9 | 168914 | 12 | AC158640 |
| 17 | 41.8 | 13.9 | 231134 | 6 | AC111046 |
| 18 | 41.6 | 13.9 | 237432 | 12 | AC123182 |

| | | | | | | |
|------|------|------|--------|----|-------------|-------------|
| c 19 | 41.6 | 13.9 | 242992 | 12 | CR792426 | CR792426 |
| c 20 | 41.4 | 13.8 | 222752 | 12 | AC094947 | AC094947 |
| c 21 | 41.4 | 13.8 | 253693 | 12 | AC131839 | AC131839 |
| c 22 | 41.4 | 13.8 | 271927 | 12 | AC098410 | AC098410 |
| c 23 | 40.8 | 13.6 | 23709 | 4 | AP004980 | AP004980 |
| c 24 | 40.8 | 13.6 | 158755 | 11 | EX255970 | EX255970 |
| c 25 | 40.8 | 13.6 | 165917 | 5 | AL356073 | AL356073 |
| c 26 | 40.4 | 13.5 | 166494 | 12 | AC135483 | AC135483 |
| c 27 | 40.4 | 13.5 | 183223 | 6 | AC142454 | AC142454 |
| c 28 | 40.4 | 13.5 | 186836 | 6 | AC139360 | AC139360 |
| c 29 | 40.4 | 13.5 | 192918 | 6 | AC126460 | AC126460 |
| c 30 | 40.4 | 13.5 | 213458 | 12 | AC123390 | AC123390 |
| c 31 | 40.4 | 13.5 | 222392 | 12 | AC095434 | AC095434 |
| c 32 | 40 | 13.3 | 110000 | 15 | AE015927_05 | AE015927_05 |
| c 33 | 40 | 13.3 | 157168 | 5 | AL137018 | AL137018 |
| c 34 | 40 | 13.3 | 160933 | 12 | AC172335 | AC172335 |
| c 35 | 40 | 13.3 | 181272 | 11 | EX548256 | EX548256 |
| c 36 | 40 | 13.3 | 192411 | 12 | AC169312 | AC169312 |
| c 37 | 39.8 | 13.3 | 4250 | 13 | AB193599 | AB193599 |
| c 38 | 39.8 | 13.3 | 157011 | 12 | AC140115 | AC140115 |
| c 39 | 39.8 | 13.3 | 161610 | 11 | AL929174 | AL929174 |
| c 40 | 39.8 | 13.3 | 185657 | 12 | AC125630 | AC125630 |
| c 41 | 39.8 | 13.3 | 207949 | 11 | EX005210 | EX005210 |
| c 42 | 39.8 | 13.3 | 208337 | 12 | AC126085 | AC126085 |
| c 43 | 39.6 | 13.2 | 46494 | 12 | AC157740 | AC157740 |
| c 44 | 39.6 | 13.2 | 110000 | 12 | AC133155_2 | AC133155_2 |
| c 45 | 39.6 | 13.2 | 123109 | 11 | EX511308 | EX511308 |
| c 46 | 39.6 | 13.2 | 154673 | 6 | AC166491 | AC166491 |
| c 47 | 39.6 | 13.2 | 165154 | 6 | AC121307 | AC121307 |
| c 48 | 39.6 | 13.2 | 193188 | 6 | AL844155 | AL844155 |
| c 49 | 39.6 | 13.2 | 200393 | 6 | AC084288 | AC084288 |
| c 50 | 39.6 | 13.2 | 229628 | 12 | EX908768 | EX908768 |
| c 51 | 39.4 | 13.1 | 48000 | 5 | AY436325 | AY436325 |
| c 52 | 39.4 | 13.1 | 60461 | 5 | AF245699 | AF245699 |
| c 53 | 39.4 | 13.1 | 110573 | 4 | AC147741 | AC147741 |
| c 54 | 39.4 | 13.1 | 159059 | 12 | AC093000 | AC093000 |
| c 55 | 39.4 | 13.1 | 160636 | 12 | AC024991 | AC024991 |
| c 56 | 39.4 | 13.1 | 163311 | 6 | AC160632 | AC160632 |
| c 57 | 39.4 | 13.1 | 204794 | 5 | AC024897 | AC024897 |
| c 58 | 39.4 | 13.1 | 223565 | 12 | AC155681 | AC155681 |
| c 59 | 39.4 | 13.1 | 255963 | 12 | AC025153 | AC025153 |
| c 60 | 39.4 | 13.1 | 270437 | 12 | CR932018 | CR932018 |
| c 61 | 39.2 | 13.1 | 161092 | 11 | CR318664 | CR318664 |
| c 62 | 39.2 | 13.1 | 189185 | 12 | AC173587 | AC173587 |
| c 63 | 39.2 | 13.1 | 202967 | 12 | AC175374 | AC175374 |
| c 64 | 39 | 13.0 | 661 | 7 | BV413550 | BV413550 |
| c 65 | 39 | 13.0 | 15500 | 13 | DQ241798 | DQ241798 |
| c 66 | 39 | 13.0 | 40852 | 13 | CEP11C1 | CEP11C1 |
| c 67 | 39 | 13.0 | 125051 | 5 | AL450164 | AL450164 |
| c 68 | 39 | 13.0 | 151319 | 5 | AC108036 | AC108036 |
| c 69 | 39 | 13.0 | 219318 | 12 | AC140987 | AC140987 |
| c 70 | 38.8 | 12.9 | 110000 | 12 | AC098516_1 | AC098516_1 |
| c 71 | 38.8 | 12.9 | 110000 | 12 | AC114829_2 | AC114829_2 |
| c 72 | 38.8 | 12.9 | 175249 | 6 | AC129603 | AC129603 |
| c 73 | 38.8 | 12.9 | 201282 | 6 | AC132450 | AC132450 |
| c 74 | 38.8 | 12.9 | 289587 | 12 | AC118337 | AC118337 |
| c 75 | 38.6 | 12.9 | 2224 | 2 | CO489965 | CO489965 |
| c 76 | 38.6 | 12.9 | 2224 | 2 | CO490293 | CO490293 |
| c 77 | 38.6 | 12.9 | 2224 | 2 | CO490299 | CO490299 |
| c 78 | 38.6 | 12.9 | 2224 | 2 | CO491306 | CO491306 |
| c 79 | 38.6 | 12.9 | 2224 | 2 | CO492635 | CO492635 |
| c 80 | 38.6 | 12.9 | 2224 | 2 | CO492712 | CO492712 |
| c 81 | 38.6 | 12.9 | 2224 | 2 | CO492723 | CO492723 |
| c 82 | 38.6 | 12.9 | 2224 | 2 | CO493179 | CO493179 |
| c 83 | 38.6 | 12.9 | 2224 | 2 | CO495812 | CO495812 |
| c 84 | 38.6 | 12.9 | 2224 | 2 | CO496154 | CO496154 |
| c 85 | 38.6 | 12.9 | 2224 | 2 | CO496160 | CO496160 |
| c 86 | 38.6 | 12.9 | 2224 | 2 | CO497174 | CO497174 |
| c 87 | 38.6 | 12.9 | 21691 | 13 | CEY17D7C | CEY17D7C |
| c 88 | 38.6 | 12.9 | 137025 | 12 | AC162158 | AC162158 |
| c 89 | 38.6 | 12.9 | 147879 | 12 | AC157887 | AC157887 |
| c 90 | 38.6 | 12.9 | 211383 | 5 | AC161278 | AC161278 |
| c 91 | 38.6 | 12.9 | 227322 | 11 | EX927261 | EX927261 |

```

c 92 38.4 12.8 110000 4 AP008209 045
c 93 38.4 12.8 110000 12 TANNIL_22
c 94 38.4 12.8 123340 4 AC151523
c 95 38.4 12.8 148894 4 AC121489
c 96 38.4 12.8 155293 12 AC154085
c 97 38.4 12.8 155633 12 AC025510
c 98 38.4 12.8 131804 5 AC020741
c 99 38.2 12.7 538 13 AY042101
c 100 38.2 12.7 45232 13 AF100657
c 101 38.2 12.7 99534 12 AC178452
c 102 38.2 12.7 127514 12 AC161985
c 103 38.2 12.7 155243 12 AC170794
c 104 38 12.7 192089 15 AF274444
c 105 38 12.7 171582 5 AC021558
c 106 38 12.7 183115 12 CT573496
c 107 38 12.7 184052 11 BX927369
c 108 38 12.7 198313 5 AC142314
c 109 38 12.7 214284 12 AC112544
c 110 38 12.7 253819 12 AC095474
c 111 37.8 12.6 12064 5 AF494057
c 112 37.8 12.6 41197 13 AF038614
c 113 37.8 12.6 67335 5 HS43683
c 114 37.8 12.6 156598 5 AC146185
c 115 37.8 12.6 168551 5 AC080007
c 116 37.8 12.6 168914 12 AC158640
c 117 37.8 12.6 173339 12 AC167778
c 118 37.8 12.6 194160 12 CR762444
c 119 37.8 12.6 194943 5 AC105252
c 120 37.8 12.6 231134 6 AC111046
c 121 37.8 12.6 245316 12 AC105887
c 122 37.6 12.5 30616 12 AC018027
c 123 37.6 12.5 101563 4 AC122543
c 124 37.6 12.5 138943 12 CR847865
c 125 37.6 12.5 163765 11 BX284684
c 126 37.6 12.5 165714 5 AC016919
c 127 37.6 12.5 171226 13 AC008144
c 128 37.6 12.5 175781 13 AC080891
c 129 37.6 12.5 176329 12 AC117968
c 130 37.6 12.5 177596 12 AC091945
c 131 37.6 12.5 178251 13 AE003737
c 132 37.6 12.5 213097 12 AC170271
c 133 37.6 12.5 225790 12 AC099213
c 134 37.6 12.5 230573 6 AC122883
c 135 37.6 12.5 244429 12 AC155061
c 136 37.4 12.5 34368 13 CEW06D12
c 137 37.4 12.5 35913 4 AC149414
c 138 37.4 12.5 63233 5 AL359504
c 139 37.4 12.5 157929 5 AL593844
c 140 37.4 12.5 160082 5 AP003695
c 141 37.4 12.5 164485 12 AC025340
c 142 37.4 12.5 164768 12 AC140766
c 143 37.4 12.5 169660 5 AP003547
c 144 37.4 12.5 189473 12 AC022950
c 145 37.4 12.5 195400 12 AC119022
c 146 37.4 12.5 200458 12 AC019237
c 147 37.4 12.5 201524 12 AC178557
c 148 37.4 12.5 203860 12 AC162535
c 149 37.4 12.5 208961 5 AC090819
c 150 37.4 12.5 210649 12 AC172121

```

ALIGNMENTS

```

RESULT 1
LOCUS AL138713 121896 bp DNA linear PRI 18-MAY-2005
DEFINITION Human DNA sequence from clone RP11-552M6 on chromosome
13q21.31-22.2 Contains part of the KLF12 gene for Kruppel-like
factor 12, complete sequence.
ACCESSION AL138713
VERSION AL138713.12 GI:14625506
KEYWORDS HTG; KLF12.

```

```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 121896)
AUTHORS Lovell, J.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
COMMENT Clone requests: clonerequest@sanger.ac.uk
On Jul 6, 2001 this sequence version replaced gi:11321828.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 13, constructed by the Sanger Centre Chromosome 13
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr13
RP11-552M6 is from the library RP11-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
FEATURES
source
1. .121896
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="13"
/map="q21.31-22.2"
/clones="RP11-552M6"
/clone_lib="RP11-11.2"
100
/misc_feature
/notes="Clone right end: RP11-474L7"
join(complement(AL139036.12:11189.11255),
complement(AL159972.8:109592.109681),
complement(AL159972.8:11438.11984),
complement(104572.104707),complement(56358.56420),
complement(6788.6945),
complement(AL160032.14:158733.168315))
/gene="KLF12"
/locus_tag="RP11-474L7.5-001"
join(complement(AL139036.12:11189.11255),
complement(AL159972.8:109592.109681),
complement(AL159972.8:11438.11984),
complement(104572.104707),complement(56358.56420),
complement(6788.6945),
complement(AL160032.14:158733.168315))
/gene="KLF12"
/product="Kruppel-like factor 12"
/notes="match: ESTs: AA007205 AA349967 AA513497 AA524929
A1373658 A1628470 AL701841 AV728478 BG834258 BG910766
F12639 R23424 R23427 R23428 T74523
match: cDNAs: AF113122 AF161471 AF330041 AJ243274 BC019680
Y14295"
join(complement(AL139036.12:11189.11221),
complement(AL159972.8:109592.109681),

```


REFERENCE

3 (bases 1 to 281593)

AUTHORS

Bovine Genome Sequencing Consortium

CONSTRM

Direct Submission

TITLE

Submitted (23-JAN-2006) Human Genomes Sequencing Center, Department

JOURNAL

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Jan 23, 2006 this sequence version replaced gi:82581398.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: FMKA

Center clone name: CH240-199K3

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 242506 bases at least Q40

Consensus quality: 245084 bases at least Q30

Consensus quality: 247259 bases at least Q20

Estimated insert size: 244496; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

* 1
 * 2000: contig of 2000 bp in length
 * 2001 2116: gap of 116 bp
 * 2117 5029: contig of 2913 bp in length
 * 5030 5079: gap of 50 bp
 * 5080 7237: contig of 2158 bp in length
 * 7238 7337: gap of unknown length
 * 7338 9400: contig of 2063 bp in length
 * 9401 9521: gap of 121 bp
 * 9522 27182: contig of 17661 bp in length
 * 27183 27232: gap of 50 bp
 * 27233 30893: contig of 3661 bp in length
 * 30894 30943: gap of 50 bp
 * 30944 39951: contig of 9008 bp in length
 * 39952 40001: gap of 50 bp
 * 40002 41556: contig of 1555 bp in length
 * 41557 41606: gap of 50 bp
 * 41607 73473: contig of 31867 bp in length
 * 73474 73523: gap of 50 bp
 * 73524 75093: contig of 1570 bp in length
 * 75094 103735: gap of 28642 bp
 * 103736 107712: contig of 3977 bp in length
 * 107713 136793: contig of 50 bp
 * 107763 136793: contig of 23031 bp in length
 * 136794 137118: gap of 325 bp
 * 137119 158502: contig of 21384 bp in length
 * 158503 158602: gap of unknown length
 * 158603 185056: contig of 26454 bp in length

* 185057 185106: gap of 50 bp
 * 185107 193804: contig of 8698 bp in length
 * 193805 193854: gap of 50 bp
 * 193855 208596: contig of 14742 bp in length
 * 208597 208646: gap of 50 bp
 * 208647 211035: contig of 2389 bp in length
 * 211036 211085: gap of 50 bp
 * 211086 224829: contig of 13744 bp in length
 * 224830 224879: gap of 50 bp
 * 224880 252287: contig of 27408 bp in length
 * 252288 252337: gap of 50 bp
 * 252338 257930: contig of 5593 bp in length
 * 257931 257980: gap of 50 bp
 * 257981 262625: contig of 4645 bp in length
 * 262626 262725: gap of unknown length
 * 262726 263878: contig of 1153 bp in length
 * 263879 264345: gap of 467 bp
 * 264346 266976: contig of 2631 bp in length
 * 266977 267076: gap of unknown length
 * 267077 268167: contig of 1091 bp in length
 * 268168 268855: gap of 688 bp
 * 268856 271722: contig of 2867 bp in length
 * 271723 271822: gap of unknown length
 * 271823 273130: contig of 1308 bp in length
 * 273131 273230: gap of unknown length
 * 273231 274398: contig of 1168 bp in length
 * 274399 274498: gap of unknown length
 * 274499 275719: contig of 1221 bp in length
 * 275720 275819: gap of unknown length
 * 275820 276893: contig of 1074 bp in length
 * 276894 276993: gap of unknown length
 * 276994 278152: contig of 1159 bp in length
 * 278153 278252: gap of unknown length
 * 278253 279307: contig of 1055 bp in length
 * 279308 279407: gap of unknown length
 * 279408 281593: contig of 2186 bp in length.

FEATURES

source

1. .281593
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="CH240-199K3"
 2001..2116
 /estimated_length=116
 5030..5079
 /estimated_length=50
 7238..7337
 /estimated_length=unknown
 9401..9521
 /estimated_length=121
 27183..27232
 /estimated_length=50
 30894..30943
 /estimated_length=50
 39952..40001
 /estimated_length=50
 41557..41606
 /estimated_length=50
 73474..73523
 /estimated_length=50
 75094..103735
 /estimated_length=28642
 107713..107762
 /estimated_length=50
 136794..137118
 /estimated_length=325

Query Match 47.7%; Score 143; DB 12; Length 281593;

Best Local Similarity 78.1%; Pred.No. 2.7e-36;

Matches 185; Conservative 0; Mismatches 50; Indels 2; Gaps 1;

64 AACAGCTAACTCCCAATGTACCTACTAGGGAGTTTATCATGAAGAAATGTTTAA 123

Qy

||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||


```

/rpt_family="Alu"
21997. .22141
/rpt_family="MaLR"
23550. .23946
/rpt_family="MaLR"
26897.26863
/rpt_family="L1"
27825. .27984
/rpt_family="B2"
29411. .29607
/rpt_family="ERVK"
31300.31876
/rpt_family="L1"
31943. .32107
/rpt_family="B4"
33292. .33385
/rpt_family="L1"
35024.35221
/rpt_family="B2"
35207. .35467
/rpt_family="L1"
35468. .35589
/rpt_family="L1"
36177.36579
/rpt_family="MaLR"
45951. .46588
/note="Sequence derived from PCR product of project DNA."
47048. .47115
/rpt_family="L1"
47115. .47807
/rpt_family="L1"
49387. .49573
/rpt_family="B2"
51674. .52329
/rpt_family="L1"
52330.53173
/rpt_family="L1"
54488. .54652
/rpt_family="B4"
54840. .54992
/rpt_family="L1"
55487.55791
/rpt_family="B4"
57343. .57541
/rpt_family="B2"
57678. .57951
/rpt_family="B4"
58729.58950
/rpt_family="MaLR"
62188. .62448
/rpt_family="B4"
63148. .63422
/rpt_family="MaLR"
65214.65481
/rpt_family="L1"
65767. .65935
/rpt_family="B4"
66245. .66353
/rpt_family="Alu"
68642.68729
/rpt_family="B4"
68920. .68997
/rpt_family="B4"
69011. .69155
/rpt_family="Alu"
69382.69672
/rpt_family="L1"
70522. .71262
/rpt_family="L1"
71271. .71381
/rpt_family="L1"

```

Query Match

17.2%; Score 51.6; DB 6; Length 175961;

```

Best Local Similarity 54.4%; Pred. No. 7.7e-06;
Matches 112; Conservative 0; Mismatches 54; Indels 8; Gaps 2;

Qy 64 AAACAGCTAACTCCCAAAATGTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAA 123
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 142554 ACATAGCCAGCTCCCAAAATGTACATAACACAGGAAATCT---ATTAAGAAATGTTT 142499
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 124 TGCTTAATTTTCAAAATAAGAAGTAACACAGAGTGTGTCACACATGCTGTGTTAATAACCCGAC 183
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 142498 TGCTTTATTTT---AGAGAAACTAACAATGTCACAAATGTCACATTCGGCTAGTAGTAGAC 142443
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 184 AACTTCAATCACTATAGTGTAGTAGAGTGCATCTTGCAGGATCCAGAGTA 237
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 142442 AACTTAAATCTCAGCAGCAATCCCTCTTAGCATTCGACAGGATGTCAGTA 142389
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 4

AC117638/c
LOCUS AC117638 184477 bp DNA linear ROD 14-AUG-2003
DEFINITION Mus musculus, clone RP23-225L9, complete sequence.
ACCESSION AC117638
VERSION AC117638.13 GI:33636787
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 184477)

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

Birren, B., Nusbaum, C. and Lander, E.
Mus musculus, clone RP23-225L9
Unpublished
2 (bases 1 to 184477)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 184477)

TITLE

JOURNAL

REFERENCE

AUTHORS

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,


```

Query Match      17.2%; Score 51.6; DB 6; Length 184477;
Best Local Similarity 64.4%; Pred. No. 7.7e-06;
Matches 112; Conservative

QY 64 AAACAGCTAACTCCCAATGTTACCTAACTAGGAGGAGTTTATCATGAAGAAATGTTTAAA 123
Db 110061 ACATAGCCACTCCCAATGTTATCATACAGGAATATCT-----ATTAAGAAATGTTTGT 110006

QY 124 TGCTTAATTTTCAAAATGAAGTAACACAGAGTGTGTCACATGCTGTTAAATACCCGAC 183
Db 110005 TGCCTTTATTTT-----AGAGAAACTAACCAAAATGTTGCAACATTCGGCTAAGTAGTAGAC 109950

QY 184 AAACCTCAATCACTATAGCTGTAGTAGAGTGCAATTCGCAAGGATCCAGAGTA 237
Db 109949 AAACCTTAATCTCAGCAGCAATCCCTCTTACGATTCGTCAGGAATGCTAGAGTA 109896

RESULT 5
AC129383
LOCUS      Rattus norvegicus clone CH230-36A10, *** SEQUENCING IN PROGRESS
DEFINITION      ***, 3 unordered pieces.
ACCESSION      AC129383
VERSION        AC129383.5 GI:24818055
KEYWORDS       HTG; HTGS-PHASE1; HTGS-DRAFT; HTGS_ENRICHED.
SOURCE         Rattus norvegicus (Norway rat)
ORGANISM       Rattus norvegicus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
               Sciurognathi; Muroidae; Muridae; Rattus.
REFERENCE
AUTHORS        Muzny, D., Maric, E., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J.,
               Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
               Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
               Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
               Blawalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
               Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
               Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Chen, Z., Chu, J.,
               Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
               Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
               Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
               Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
               Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
               Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
               Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
               Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
               Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
               Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
               Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
               Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
               Hollins, B., Howells, S., Hulyk, S., Hume, J., Idiebird, D., Jackson, A.,
               Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
               Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
               Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
               Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
               Lorensuhera, L., Louissegh, H., Lozada, R.J., Lu, X., Ma, J.,
               Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A.,
               Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
               Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
               Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
               Morgan, M., Morris, K., Morris, S., Munday, M., Murphy, M., Nair, L.,
               Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
               Nwakeleneh, O., Okwunonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
               Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
               Plopper, P., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
               Puazo, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M.A., Reigh, R.,
               Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
               Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
               Sanders, W., Savary, G., Scherer, S., Scott, G., Shateman, S., Shen, H.,
               Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
               Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, J.,
               Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
               Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,

```

```

Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, V., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 246893)
Worley, K.C.
Direct Submission
Submitted (29-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 246893)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 9, 2002 this sequence version replaced gi:23096176.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
----- Project name: GEEP
Center clone name: CH230-36A10
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 225900 bases at least Q40
Consensus quality: 229196 bases at least Q30
Consensus quality: 231050 bases at least Q20
Estimated insert size: 230916; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 243625: contig of 243625 bp in length
* 243626 243725: gap of unknown length
* 243726 244798: contig of 1073 bp in length
* 244799 244898: gap of unknown length
* 244899 246893: contig of 1995 bp in length.
* Location/Qualifiers
* 1..246893
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clones="CH230-36A10"
* 1..1304
* /note="wgs_contig"
FEATURES
source
misc_feature

```

```

misc_feature 3950..5173
    note="wgs_contig"
misc_feature 140573..142683
    note="wgs_contig"
misc_feature complement(233360..234110)
    note="clone_boundary"
    clone_end:17
    site:ECORI
end_sequence:BH299246"
misc_feature 234187..235820
    note="wgs_end_extension"
    clone_end:17"
misc_feature 242264..243625
    note="wgs_end_extension"
    clone_end:17"
gap 243626..243725
    /estimated_length=unknown
gap 244799..244898
    /estimated_length=unknown
ORIGIN
Query Match 14.7%; Score 44.2; DB 12; Length 246893;
Best Local Similarity 61.8%; Pred. No. 0.0022;
Matches 107; Conservative 0; Mismatches 58; Indels 8; Gaps 2;
Qy 64 AAACGCTAATCTCCCAATGTGACCTAATAGGAGGTTTATCATGAAGAATGTTTAAA 123
Db 135556 ACACAGCCCATCTCCCAATGTATATACAGGAAA-----TTCATGAGGAATGTTTCA 135611
Qy 124 TGCTTAATTTCAATATAGAGTACACAGAGTGTTCACATCTGTTAAATACCCGAC 183
Db 135612 TGCTTTGTTTAGAA-----AAACTAACCAAGTGTGCAACACTTTGTTAAGTCGTAGAC 135667
Qy 184 AAATCTCAATCACTAGTCTAGTAGAGTGCATCTGCAAGGATCCACAGAGT 236
Db 135668 AAACGGAATCTCTAGCAATCTCTTTTACCATTCTGAGAAATCTGGACT 135720

RESULT 6
AC130228/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-53K23, *** SEQUENCING IN PROGRESS
***, 2 unordered pieces.
ACCESSION AC130228
VERSION AC130228.4 GI:24635276
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 314404)
Muzny D.Marie, Metzker M.Lee, Abramson S., Adams C., Alder J.,
Allen C., Allen H., Albrooks S., Amin A., Anguiano D.,
Anyalebechi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,
Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,
Biswal K., Blair J., Blankenburg K., Blyth P., Brown M.,
Bryant N., Buhay C., Burch P., Burrell K., Calderon E.,
Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,
Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,
Cleaveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L.,
Davila M.L., Davis C., Davy-Carroll L., De Anda C., Dederich D.,
Delgado O., Denson S., Deramo C., Ding Y., Dinh H., Divya K.,
Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K.,
Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G.,
Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,
Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M.,
Gebregeorgis E., Geer K., Gill R., Grady M., Guerra W., Guevara W.,
Gunaratne P., Haaland W., Hamil C., Hamilton C., Hamilton K.,
Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J.,
Hernandez R., Hines S., Hladun S.L., Hodgson A., Hogue M.,
Hollins B., Howells S., Hulyk S., Hume J., Idlebird D., Jackson A.,
Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A.,

```

```

Karpthy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensuhewa, L., Loulseged, H., Lozado, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McJeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokelimeh, O., Okwuonu, G., Olarnpungoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfanckoch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smajs, D.,
Sneid, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
Valas R., Vera, V., Villanasa, D., Waldron, L., Walker, B., Wang, J.,
Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (Bases 1 to 314404)
Worley, K.C.
Direct Submission
Submitted (09-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (Bases 1 to 314404)
Rat Genome Sequencing Consortium.
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 6, 2002 this sequence version replaced gi:23264801.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KCIZ
Center clone name: CH230-53K23
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 242417 bases at least Q40
Consensus quality: 244920 bases at least Q30
Consensus quality: 246473 bases at least Q20
Estimated insert size: 246166; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

```

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 312945: contig of 312945 bp in length
 * 312946 313045: gap of unknown length
 * 313046 314404: contig of 1359 bp in length.

FEATURES

```

source
    Location/Qualifiers
      1..314404
        /organism="Rattus norvegicus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10116"
        /clone="CH230-53K23"
      1..1916
        /note="wgs end extension"
        /clone_end:Sp6"
        2588..4415
        /note="wgs end extension"
        /clone_end:Sp6"
        complement(5720..6427)
        /note="clone_boundary"
        /clone_end:Sp6"
        site:
          end sequence:BH323709"
          242173..243018
          /note="clone_boundary"
          /clone_end:T7"
        site:
          end sequence:BH323707"
          243109..244516
          /note="wgs end extension"
          /clone_end:T7"
          245081..246984
          /note="wgs end extension"
          /clone_end:T7"
          310363..312945
          /note="wgs end extension"
          /clone_end:T7"
          312946..313045
          /estimated_length=unknown

```

ORIGIN

```

Query Match      14.7%; Score 44.2; DB 12; Length 314404;
Best Local Similarity 61.8%; Pred. No. 0.0022;
Matches 107; Conservative 0; Mismatches 56; Indels 8; Gaps 2;

Qy 64 AAACAGCTAACTCCCAATTTACTACTAGGGGAGTTTATCATGAGAAATGTTTAA 123
Db 41164 ACACAGCCATCCCAATGTATATACAGGAAA---TTCATGAGGAATGTTTCA 41109

Qy 124 TGCTTAATTTCAATAGAGTAGTACACAGAGTGTTCACATGCTGTTAATATACCCGAC 183
Db 41108 TGCTTTGTTTGTAGAA---AAACTAACCAAGTGTGCAACACTTTTGTAACTCGCTAGAC 41053

Qy 184 AAATCTCAATCATATAGCTAGTAGTGCATTCGCAAGGATCCAGAGT 236
Db 41052 AAACGGAATCTGTAGCAATTCCTTTTACCATTCGCAAAATATCTGGACT 41000

```

RESULT 7

```

AC153067/c
LOCUS AC153067 218961 bp DNA linear HTG 01-JUL-2005
DEFINITION Bos taurus clone CH240-1302, *** SEQUENCING IN PROGRESS ***, 31
unordered pieces.
ACCESSION AC153067
VERSION AC153067.3 GI:68226955
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus

```

REFERENCE AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
 1 (bases 1 to 218961)
 Muzny,D,Marie,, Metzker,M, Lee,, Abramzon,S,, Adams,C,, Alder,J,, Allen,C,, Allen,H,, Alsbrooks,S,, Amin,A,, Anguiano,D,, Anyalebechi,V,, Aoyagi,A,, Ayodeji,M,, Baca,E,, Baden,H,, Baldwin,D,, Bandaranaike,D,, Barber,M,, Barnstead,M,, Benahmed,F,, Biswal,K,, Blair,J,, Blankenburg,K,, Blyth,P,, Brown,M,, Bryant,N,, Buhay,C,, Burch,P,, Burrell,K,, Calderon,E,, Cardenas,V,, Carter,K,, Cavazos,I,, Ceasar,H,, Center,A,, Chacko,J,, Chavez,D,, Chen,G,, Chen,R,, Chen,Y,, Chen,Z,, Chu,J,, Cleveland,C,, Cockrell,R,, Cox,C,, Coyle,M,, Cree,A,, D'Souza,L,, Davila,M,L,, Davis,C,, Davy-Carroll,L,, De Anda,C,, Dederich,D,, Delgado,O,, Denson,S,, Deramo,C,, Ding,Y,, Dinh,H,, Divya,K,, Draper,H,, Dugan-Rocha,S,, Dunn,A,, Durbin,K,, Duval,B,, Eaves,K,, Egan,A,, Escotto,M,, Eugene,C,, Evans,C,A,, Falls,T,, Fan,G,, Fernandez,S,, Finley,M,, Flagg,N,, Forbes,L,, Foster,M,, Foster,P,, Fraser,C,M,, Gabisi,A,, Ganta,R,, Garcia,A,, Garner,T,, Garza,M,, Gebregeorgis,E,, Geer,K,, Gill,R,, Grady,M,, Guerra,W,, Guevara,W,, Gunaratne,P,, Haaland,W,, Hamil,C,, Hamilton,C,, Hamilton,K,, Harvey,Y,, Havlak,P,, Hawes,A,, Henderson,N,, Hernandez,J,, Hernandez,R,, Hines,S,, Hladun,S,L,, Hodgson,A,, Hogues,M,, Hollins,B,, Howells,S,, Hulyk,S,, Hume,J,, Idlebird,D,, Jackson,A,, Jackson,L,, Jacob,L,, Jiang,H,, Johnson,B,, Johnson,R,, Jolivet,A,, Karpathy,S,, Kelly,S,, Kelly,S,, Khan,Z,, King,L,, Kovat,C,, Kowis,C,, Kraft,C,L,, Lebow,H,, Levan,J,, Lewis,L,, Li,Z,, Liu,J,, Liu,J,, Liu,W,, Liu,Y,, London,P,, Longacre,S,, Lopez,J,, Lorensuhewa,L,, Loulseghe,H,, Lozado,R,J,, Lu,X,, Ma,J,, Maheshwari,M,, Mahindartne,M,, Mahmoud,M,, Malloy,K,, Mangum,A,, Mangun,B,, Mapua,P,, Martin,K,, Martin,R,, Martinez,E,, Mawhney,S,, McLeod,M,P,, McNeill,T,Z,, Meenen,E,, Milosavljevic,A,, Miner,G,, Minja,E,, Montemayor,J,, Moore,S,, Morgan,M,, Morris,K,, Morris,S,, Munidasa,M,, Murphy,M,, Nair,L,, Nankervis,C,, Neal,D,, Newton,N,, Nguyen,N,, Norris,S,, Nwaokemeleh,O,, Okwuonu,G,, Olarnpunsagoon,A,, Pal,S,, Parks,K,, Pasternak,S,, Paul,H,, Perez,A,, Perez,L,, Primus,E,, Pu,L,, L., Plopper,F,, Poindexter,A,, Popovic,D,, Reeves,K,, Regier,M,A,, Reigh,R,, Puazo,M,, Quiroz,J,, Rachlin,E,, Reeves,K,, Richards,S,, Riggs,F,, Reilly,B,, Reilly,M,, Ren,Y,, Reuter,M,, Richards,S,, Riggs,F,, Rives,C,, Rodkey,T,, Rojas,A,, Rose,M,, Rose,R,, Ruiz,S,J,, Sanders,W,, Savery,G,, Scherer,S,, Scott,G,, Shatsman,S,, Shen,H,, Shetty,J,, Shvartsbeyn,A,, Sisson,I,, Sitter,C,D,, Smajs,D,, Sneed,A,, Sodergren,E,, Song,X,-Z,, Sorelle,R,, Sosa,J,, Steimle,M,, Strong,R,, Sutton,A,, Svatek,A,, Taber,P,, Taylor,C,, Taylor,T,, Thomas,N,, Thomas,S,, Tingey,A,, Trejos,Z,, Usmani,K,, Valas,R,, Vera,V,, Villaseana,D,, Waldron,L,, Walker,B,, Wang,J,, Wang,Q,, Wang,S,, Warren,J,, Warren,R,, Wei,X,, White,F,, Williams,G,, Willson,R,, Wleczyk,R,, Wooden,H,, Worley,K,, Wright,D,, Wright,R,, Wu,J,, Yakub,S,, Yen,J,, Yoon,L,, Yoon,V,, Yu,F,, Zhang,J,, Zhou,J,, Zhou,X,, Zhao,S,, Dunn,D,, von Niederhausern,A,, Weiss,R,, Smith,D,R,, Holt,R,A,, Smith,H,O,, Weinstock,G. and Gibbs,R.A.

TITLE

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Unpublished
 2 (bases 1 to 218961)
 Direct Submission
 Worley,K.C.

Submitted (01-DEC-2004) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 218961)
 Direct Submission
 Cow Genome Sequencing Consortium.

Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Jun 26, 2005 this sequence version replaced gi:58037848. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FAXO
Center clone name: CH240-13J2
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 203243 bases at least Q40
Consensus quality: 206548 bases at least Q30
Consensus quality: 209635 bases at least Q20
Estimated insert size: 212382; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation table.

----- NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
----- NOTE: This is a 'working draft' sequence. It currently consists of 31 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 12238: contig of 12238 bp in length
12239 12515: gap of 277 bp
12516 14308: contig of 1793 bp in length
14309 14408: gap of unknown length
14409 23603: contig of 9195 bp in length
23604 23653: gap of 50 bp
23654 28728: contig of 5075 bp in length
28729 38011: contig of 9233 bp in length
38012 38061: gap of 50 bp
38062 41982: contig of 3921 bp in length
41983 42032: gap of 50 bp
42033 47531: contig of 5499 bp in length
47532 47581: gap of 50 bp
47582 82655: contig of 35074 bp in length
82656 82705: gap of 50 bp
82706 84911: contig of 2206 bp in length
84912 85011: gap of unknown length
85012 86758: contig of 1747 bp in length
86759 86809: gap of 50 bp
86809 103654: contig of 16846 bp in length
103655 103704: gap of 50 bp
103705 108056: contig of 4352 bp in length
108057 108156: gap of unknown length
108157 110213: contig of 2057 bp in length
110214 114233: gap of 1210 bp
11424 113681: contig of 2258 bp in length
113682 113781: gap of unknown length
113782 115505: contig of 1724 bp in length
115506 115506: gap of 50 bp
115556 122394: contig of 6839 bp in length
122395 122494: gap of unknown length
122495 124304: contig of 1810 bp in length
124305 124354: gap of 50 bp
124355 144012: contig of 19658 bp in length
144013 144924: gap of 912 bp
144925 150751: contig of 5827 bp in length
150752 150801: gap of 50 bp
150802 154839: contig of 4038 bp in length
154840 154889: gap of 50 bp

* 154890 161537: contig of 6648 bp in length
* 161538 161587: gap of 50 bp
* 161588 171569: contig of 9982 bp in length
* 171570 171619: gap of 50 bp
* 171620 181844: contig of 10225 bp in length
* 181845 181894: gap of 50 bp
* 181895 184328: contig of 2434 bp in length
* 184329 184901: gap of 573 bp
* 184902 199708: contig of 14807 bp in length
* 199709 199758: gap of 50 bp
* 199759 210817: contig of 11059 bp in length
* 210818 210917: gap of unknown length
* 210918 211927: contig of 1010 bp in length
* 211928 212027: gap of unknown length
* 212028 213436: contig of 1409 bp in length
* 213437 213536: gap of unknown length
* 213537 214731: contig of 1195 bp in length
* 214732 214831: gap of unknown length
* 214832 217464: contig of 2533 bp in length
* 217465 217465: gap of unknown length
* 217465 218961: contig of 1497 bp in length.

FEATURES

Location/Qualifiers
1. 218961
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-13J2"
12239..12515
/estimated length=277
14309..14408
/estimated length=unknown
23604..23653
/estimated length=50
28729..28778
/estimated length=50
38012..38061
/estimated length=50
41983..42032
/estimated length=50
47532..47581
/estimated length=50
82656..82705
/estimated length=50
84912..85011
/estimated length=unknown
86759..86808

Query Match 14.6%; Score 43.8; DB 12; Length 218961;
Best Local Similarity 51.8%; Pred. No. 0.0031;
Matches 99; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 5 TTGACAGCTAAATGCTAAGTGACTACGATGAAGACTTGAATATCTTCATTTGAATA 64
Db 40349 TTGACATATAAATGACAAAGCAATAATATGTTGAAATATATTTATATTTGCATTT 40290
Qy 65 AACAGCTAACTCCCAATTTGACCTAATGAGGAGTTTATCATGAAGAAATGTTTAAAT 124
Db 40289 ATAAAGGCTCTTCTGATATATTTCTCACTATTTATGTTTTATAGTGAAGTAGAATT 40230
Qy 125 GCTTAATTTCAATAAAGAGTAACCAAGAGTGTGTGCAACATGCTGTGTTAAATACCCGACA 184
Db 40229 GCTTATTATAGTAAGATTATATAACATATTGTTGCAAAAGCAATTCATTTCAAACCCAC 40170
Qy 185 AACTTCAATCA 195
Db 40169 AAGTTCAATCA 40159

RESULT 8
DMU11584
LOCUS DMU11584 4601 bp DNA linear INV 23-JUL-1994
DEFINITION Drosophila melanogaster Oregon-R mitochondrial A+T region.
ACCESSION U11584

```

VERSION      U11584.1  GI:508826
KEYWORDS     mitochondrial DNA; A+T region; tandem repeats.
SOURCE       Drosophila melanogaster (fruit fly)
ORGANISM     Drosophila melanogaster
              Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
              Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
              Ephydroidea; Drosophilidae; Drosophila.
REFERENCE    1 (bases 1 to 4601)
AUTHORS      Lewis,D.L., Farr,C.L., Farquhar,A.L. and Kaguni,L.S.
TITLE        Sequence, organization, and evolution of the A+T region of
              Drosophila melanogaster mitochondrial DNA
JOURNAL      Mol. Biol. Evol. 11 (3), 523-538 (1994)
PUBMED       8015445
REFERENCE    2 (bases 1 to 4601)
AUTHORS      Kaguni,L.S.
TITLE        Direct Submission
JOURNAL      Submitted (28-JUN-1994) Laurie S. Kaguni Ph.D, Dept. of
              Biochemistry, Michigan State University, East Lansing, MI,
              48824-1318, USA
FEATURES     source
              1..4601
               /organism="Drosophila melanogaster"
               /organella="mitochondrion"
               /mol_type="genomic DNA"
               /strain="Oregon-R"
               /db_xref="taxon:7227"
               /dev_stage="embryo"
               650..1022
               /note="repeat I-A"
               /rpt_type=tandem
               1023..1360
               /note="repeat I-B1"
               /rpt_type=tandem
               1361..1705
               /note="repeat I-C/A"
               /rpt_type=tandem
               1706..2043
               /note="repeat I-B2"
               /rpt_type=tandem
               2044..2388
               /note="repeat I-C"
               /rpt_type=tandem
               2491..2511
               /note="deoxythymidylate stretch"
               2512..2648
               /rpt_type=tandem
               2649..3112
               /note="repeat II-A"
               /rpt_type=tandem
               3113..3576
               /note="repeat II-B1"
               /rpt_type=tandem
               3577..4040
               /note="repeat II-B2"
               /rpt_type=tandem
               4041..4504
               /note="repeat II-C"
               /rpt_type=tandem
               complement(4565..4585)
               /note="deoxythymidylate stretch"
ORIGIN
Query Match      14.5%; Score 43.6; DB 13; Length 4601;
Best Local Similarity 58.5%; Pred. No. 0.0038;
Matches 76; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
29 TACGATGAAGACTTGAATATCTTCATTTTGAATAAACACGCTAACCCCAATTTGACC 88
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1494 TATATATTAACCTTAAACAAAATATTTTTTTTAAAAAAAATATTTTATTAATATACT 1553
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
89 TAACAGGGAGTTTATCATCAAGAAATGTTTAATGCTTAATTTTCAATAAAGAGTAA 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1554 TAATAAACTATTTTATTAATAAATATTTTATTAATAAATAATATTTAATAATAATA 1613
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Qy      149 CCAGAGTGTT 158
Db      1614 GAAATATTTT 1623

RESULT 9
DMU37541
LOCUS     19517 bp      DNA      circular INV 02-MAR-2001
DEFINITION Drosophila melanogaster complete mitochondrial genome.
ACCESSION U37541
VERSION   U37541.1  GI:1166529
KEYWORDS  .
SOURCE    mitochondrion Drosophila melanogaster (fruit fly)
ORGANISM  Drosophila melanogaster
              Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
              Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
              Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 12511 to 12682)
AUTHORS   Clary,D.O., Goddard,J.M., Martin,S.C., Fauron,C.M. and
              Wolstenholme,D.R.
TITLE     Drosophila mitochondrial DNA: a novel gene order
JOURNAL   Nucleic Acids Res. 10 (21), 6619-6637 (1982)
PUBMED    6294611
REFERENCE 2 (bases 5269 to 5695)
AUTHORS   Clary,D.O., Wahleithner,J.A. and Wolstenholme,D.R.
TITLE     Transfer RNA genes in Drosophila mitochondrial DNA: related 5'
              flanking sequences and comparisons to mammalian mitochondrial tRNA
              genes
JOURNAL   Nucleic Acids Res. 11 (8), 2411-2425 (1983)
PUBMED    6304652
REFERENCE 3 (bases 404 to 5272)
AUTHORS   de Bruijn,M.H.
TITLE     Drosophila melanogaster mitochondrial DNA, a novel organization and
              genetic code
JOURNAL   Nature 304 (5923), 234-241 (1983)
PUBMED    6408489
REFERENCE 4 (bases 804 to 1778)
AUTHORS   Satta,Y., Ishiwa,H. and Chigusa,S.I.
TITLE     Analysis of nucleotide substitutions of mitochondrial DNAs in
              Drosophila melanogaster and its sibling species
JOURNAL   Mol. Biol. Evol. 4 (6), 638-650 (1987)
PUBMED    2832697
REFERENCE 5 (bases 5268 to 13619)
AUTHORS   Garesse,R.
TITLE     Drosophila melanogaster mitochondrial DNA: gene organization and
              evolutionary considerations
JOURNAL   Genetics 118 (4), 649-663 (1988)
PUBMED    3130291
REFERENCE 6 (bases 441 to 2967)
AUTHORS   Satta,Y. and Takahata,N.
TITLE     Evolution of Drosophila mitochondrial DNA and the history of the
              melanogaster subgroup
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9558-9562 (1990)
PUBMED    2124697
REFERENCE 7 (bases 14215 to 14512)
AUTHORS   Ballard,J.W., Olsen,G.J., Faith,D.P., Odgers,W.A., Rowell,D.M. and
              Atkinson,P.W.
TITLE     Evidence from 12S ribosomal RNA sequences that onychophorans are
              modified arthropods
JOURNAL   Science 258 (5086), 1345-1348 (1992)
PUBMED    1455227
REFERENCE 8 (bases 14917 to 19517)
AUTHORS   Lewis,D.L., Farr,C.L., Farquhar,A.L. and Kaguni,L.S.
TITLE     Sequence, organization, and evolution of the A+T region of
              Drosophila melanogaster mitochondrial DNA
JOURNAL   Mol. Biol. Evol. 11 (3), 523-538 (1994)
PUBMED    8015445
REFERENCE 9 (bases 1 to 408; 13319 to 19517)
AUTHORS   Lewis,D.L., Farr,C.L. and Kaguni,L.S.
TITLE     Drosophila melanogaster mitochondrial DNA: completion of the
              nucleotide sequence and evolutionary comparisons
JOURNAL   Insect Mol. Biol. 4 (4), 263-278 (1995)

```



```
Contact: nisc_mouse@nhgri.nih.gov
----- Project Information
Center project name: qc
Center clone name: 343A21
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 179429 bases at least Q40
Consensus quality: 185220 bases at least Q30
Consensus quality: 188059 bases at least Q20
Insert size: 196000; agarose-fp
Insert size: 192940; sum-of-contigs
Quality coverage: 4.03x in Q20 bases; agarose-fp
Quality coverage: 4.09x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 1293: contig of 1293 bp in length
* 1294: gap of unknown length
* 1394: 2711: contig of 1318 bp in length
* 2712: 2811: gap of unknown length
* 2812: 3945: contig of 1134 bp in length
* 3946: 4045: gap of unknown length
* 4046: 5332: contig of 1287 bp in length
* 5333: 5432: gap of unknown length
* 5433: 7745: contig of 2313 bp in length
* 7746: 7845: gap of unknown length
* 7846: 10021: contig of 2176 bp in length
* 10022: 10122: gap of unknown length
* 10122: 13447: contig of 3326 bp in length
* 13448: 13547: gap of unknown length
* 13548: 16926: contig of 3379 bp in length
* 16927: 17026: gap of unknown length
* 17027: 21840: contig of 4814 bp in length
* 21841: 21940: gap of unknown length
* 21941: 26810: contig of 4870 bp in length
* 26811: 26911: gap of unknown length
* 26911: 31227: contig of 4317 bp in length
* 31228: 31327: gap of unknown length
* 31328: 36553: contig of 5226 bp in length
* 36554: 36654: gap of unknown length
* 36654: 42253: contig of 5600 bp in length
* 42254: 42354: gap of unknown length
* 42354: 46755: contig of 4402 bp in length
* 46756: 46856: gap of unknown length
* 46856: 51327: contig of 4471 bp in length
* 51327: 51427: gap of unknown length
* 51427: 55184: contig of 3758 bp in length
* 55185: 55284: gap of unknown length
* 55285: 60367: contig of 5083 bp in length
* 60368: 60467: gap of unknown length
* 60468: 65557: contig of 5090 bp in length
* 65558: 65658: gap of unknown length
* 65658: 70384: contig of 4727 bp in length
* 70385: 70484: gap of unknown length
* 70485: 75501: contig of 5017 bp in length
* 75502: 75602: gap of unknown length
* 75602: 81571: contig of 5969 bp in length
* 81571: 81671: gap of unknown length
* 81671: 86308: contig of 4638 bp in length
* 86309: 86408: gap of unknown length
* 86409: 92383: contig of 5975 bp in length
* 92384: 92483: gap of unknown length
* 92484: 97120: contig of 4637 bp in length
* 97121: 97220: gap of unknown length
* 97221: 104915: contig of 7695 bp in length
*
* 104916: 105015: gap of unknown length
* 105016: 112419: contig of 7404 bp in length
* 112420: 120438: contig of 7919 bp in length
* 120439: 120538: gap of unknown length
* 120539: 128349: contig of 7811 bp in length
* 128350: 128449: gap of unknown length
* 128450: 144135: contig of 15686 bp in length
* 144136: 144236: gap of unknown length
* 144236: 155576: contig of 11341 bp in length
* 155577: 155676: gap of unknown length
* 155677: 167613: contig of 11943 bp in length
* 167620: 167719: gap of unknown length
* 167720: 180079: contig of 12360 bp in length
* 180080: 180179: gap of unknown length
* 180180: 196140: contig of 15961 bp in length.
*
FEATURES
    source
    Location/Qualifiers
        1..196140
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /strain="C57BL6/J"
            /db_xref="taxon:10090"
            /chromosome="12"
            /clone="RP23-343A21"
            /clone_lib="RPCI mouse BAC library 23"
        1..1293
            /note="assembly_fragment"
        1294..1393
            /estimated_length=unknown
        1394..2711
            /note="assembly_fragment"
        2712..2811
            /estimated_length=unknown
        2812..3945
            /note="assembly_fragment"
        3946..4045
            /estimated_length=unknown
        4046..5332
            /note="assembly_fragment"
        5333..5432
            /estimated_length=unknown
        5433..7745
            /note="assembly_fragment"
        7746..7845
            /estimated_length=unknown
        7846..10021
            /note="assembly_fragment"
        clone_end:SP6
        vector_slide:right
        10022..10121
            /estimated_length=unknown
        10122..13447
            /note="assembly_fragment"
        13448..13547
            /estimated_length=unknown
        13548..16926
            /note="assembly_fragment"
        16927..17026
            /estimated_length=unknown
        17027..21840
            /note="assembly_fragment"
        21841..21940
            /estimated_length=unknown
        21941..26810
            /note="assembly_fragment"
        26811..26910
            /estimated_length=unknown
        26911..31227
            /note="assembly_fragment"
        31228..31327
            /estimated_length=unknown
        31328..36553
            /note="assembly_fragment"
```

```

gap          36554..36653
/estimated_length=unknown
misc_feature 36654..42253
/estimated_length=unknown
gap          42254..42353
/estimated_length=unknown
misc_feature 42354..46755
/estimated_length=unknown
gap          46756..46855
/estimated_length=unknown
misc_feature 46856..51326
/estimated_length=unknown
gap          51327..51426
/estimated_length=unknown
misc_feature 51427..55184
/estimated_length=unknown
gap          55185..55284
/estimated_length=unknown
misc_feature 55285..60367
/estimated_length=unknown
gap          60368..60467
/estimated_length=unknown
misc_feature 60468..65557
/estimated_length=unknown

Query Match      14.3%; Score 43; DB 12; Length 196140;
Best Local Similarity 56.8%; Pred. No. 0.0057;
Matches 79; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 111 AGAATGTTTAAAGCTTAATTTTCAATAGAGAGTACACGAGTGTTCACACATGCTGT 170
Db 128246 ACAATGATCATGTCGACTAATATTATTATAAGGAAAAAGCATGCTGAAGGAAGAGTA 128187

QY 171 TAAATAACCCGACAAAACCTTCAATCACTACTGTAGTAGAGTGCATTTCTGCAAGGATCC 230
Db 128186 AAGTTAGCTTAAGATTTTAAATAATATAGTTGAGTAGTGTATGAGCCACAGGACAC 128127

QY 231 CAGAGTACACAGTATTTTG 249
Db 128126 AAGACTATCTCTTAATTTG 128108

```

```

RESULT 12
LOCUS AC137708 154779 bp DNA linear ROD 16-DEC-2005
DEFINITION Mus musculus chromosome 3, clone RP24-277H8, complete sequence.
ACCESSION AC137708
VERSION AC137708.16 GI:83699489
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 154779)
Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus chromosome 3, clone RP24-277H8
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 154779)

```

```

REFERENCE 2 (bases 1 to 154779)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
TITLE Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
JOURNAL Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collumore,A.,
AUTHORS Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gordon,L., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
Minghova,T., Meneus,L., Meneus,L., Nelson,J., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Riese,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodores,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (29-JUL-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 154779)

```

```

Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodores,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (28-NOV-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 154779)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collumore,A., Cook,A., Cooke,P., Corum,B.,
Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Meneus,L., Mihova,T., Minghova,T.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Riese,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodores,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (07-JUL-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 154779)

```

```

TITLE Direct Submission
JOURNAL Submitted (07-JUL-2005) Broad Institute of MIT and Harvard, 320
REFERENCE Charles Street, Cambridge, MA 02141, USA
AUTHORS 4 (bases 1 to 154779)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collumore,A., Cook,A., Cooke,P., Corum,B.,
Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Meneus,L., Mihova,T., Minghova,T.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Riese,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodores,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (29-JUL-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 154779)

```

```

TITLE Direct Submission
JOURNAL Submitted (29-JUL-2005) Broad Institute of MIT and Harvard, 320
REFERENCE Charles Street, Cambridge, MA 02141, USA
AUTHORS 5 (bases 1 to 154779)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collumore,A., Cook,A., Cooke,P., Corum,B.,
Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Meneus,L., Mihova,T., Minghova,T.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Riese,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodores,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (29-JUL-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 154779)

```



```

Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (07-NOV-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 212071)
REFERENCE
AUTHORS
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,I., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collimore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
McCarthy,M., Macdonald,P., Major,J., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (19-JUL-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 212071)
REFERENCE
AUTHORS
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,I., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collimore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-SEP-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
On Sep 16, 2005 this sequence version replaced gi:28951332.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Broad Institute of MIT and Harvard
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: L19190
Center clone name: 176_K_16
-----
Some of the sequence contained within base pairs 104864 to the end

```

```

of the clone was stolen from accession AC121861.
FEATURES
source
1. .212071
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone="RP23-176K16"
/clone_lib="RPCI-23 Female Mouse BAC"
1720. .1766
/rpt_family="(TG)n"
repeat_region
1949. .2094
/rpt_family="B1_Mus2"
repeat_region
2101. .2138
/rpt_family="(CAAAC)n"
repeat_region
complement(7429. .7512)
/rpt_family="B3A"
repeat_region
9625. .9653
/rpt_family="(TTTG)n"
repeat_region
10718. .11779
/rpt_family="Lx2"
repeat_region
11780. .11986
/rpt_family="Lx2"
repeat_region
13111. .13142
/rpt_family="(CA)n"
repeat_region
14477. .14497
/rpt_family="(T)n"
repeat_region
14665. .14760
/rpt_family="L2"
repeat_region
15019. .15229
/rpt_family="B3"
repeat_region
15257. .15431
/rpt_family="B2_Mm2"
repeat_region
15642. .15662
/rpt_family="AT_rich"
repeat_region
15888. .15931
/rpt_family="(CTCA)n"
repeat_region
17432. .17564
/rpt_family="B1_Mus1"
repeat_region
17620. .17695
/rpt_family="B3A"
repeat_region
complement(20050. .20206)
/rpt_family="B3"
repeat_region
21499. .21536
/rpt_family="(TC)n"
repeat_region
21537. .21571
/rpt_family="(CA)n"
repeat_region
22607. .22642
/rpt_family="A-rich"
repeat_region
24493. .24520
/rpt_family="AT_rich"
repeat_region
26140. .26313
/rpt_family="B3"
repeat_region
26457. .26477
/rpt_family="AT_rich"
repeat_region
26772. .26858
/rpt_family="RSINE1"
repeat_region
27249. .27296
/rpt_family="(TAAA)n"
repeat_region
29068. .30649
/rpt_family="L1VL1"
repeat_region
30640. .31545
/rpt_family="L1_Mm"
repeat_region
31549. .31940
/rpt_family="L1_Mus3"
repeat_region
33489. .33509
/rpt_family="AT_rich"
repeat_region
complement(34372. .34508)
/rpt_family="B1_Mm"
repeat_region
complement(34528. .34616)
/rpt_family="B3"
repeat_region
35284. .35307

```


Center project name: L26568
Center clone name: 194_N_4

```

FEATURES
    source      Location/Qualifiers
1..223110     /organism="Mus musculus"
               /mol_type="Genomic DNA"
               /db_xref="taxon:10090"
               /chromosome="1"
               /map="1"
               /clone="RP24-194N4"
               /clone_lib="RPCI-24 Male Mouse BAC"
15..106      /rpt_family="B4A"
107..380     /rpt_family="L1_MM"
               complement(317..335)
               /notes="single clone coverage"
381..576     /rpt_family="B4A"
               complement(521..638)
               /notes="single clone coverage"
634..678     /rpt_family="(TC)n"
678..714     /rpt_family="(TG)n"
               complement(733..929)
               /rpt_family="B4A"
1751..1921   /rpt_family="B2_Mm2"
               complement(3408..3513)
               /rpt_family="RLTR11A"
3514..3545   /rpt_family="(TA)n"
               complement(3546..3935)
               /rpt_family="RLTR11A"
4607..4643   /rpt_family="(TTTTTG)n"
               complement(4648..4794)
               /rpt_family="B1_MM"
5568..5667   /rpt_family="RSINE1"
5894..6046   /rpt_family="B1F"
6066..6199   /rpt_family="PB1D10"
               complement(7539..7605)
               /rpt_family="ID3"
               complement(7628..7794)
               /rpt_family="B3"
8256..8341   /rpt_family="TD_B1"
8342..8458   /rpt_family="PB1D7"
9129..9318   /rpt_family="ID_B1"
               complement(9838..9986)
               /rpt_family="MIR"
               complement(10067..10165)
               /rpt_family="PB1D10"
12439..12689 /rpt_family="TD_B1"
12850..13097 /rpt_family="B4A"
13457..13477 /rpt_family="(CAGGG)n"
               complement(13690..13791)
               /rpt_family="PB1D7"
               complement(14756..14867)
               /rpt_family="PB1D7"
               complement(15026..15106)
               /rpt_family="L2"
               complement(15324..15570)

```

```

/rpt_family="L2"
16162..16187 /rpt_family="(CA)n"
16187..16248 /rpt_family="(GA)n"
               complement(17199..17340)
               /rpt_family="B1_MM"
               complement(17372..17509)
               /rpt_family="B1_MM"
17923..18588 /rpt_family="Lx2B"
               complement(19650..19754)
               /rpt_family="PB1D7"
20478..20611 /rpt_family="L2"
               complement(20772..20914)
               /rpt_family="B1_MM"
21465..21881 /rpt_family="MTC"
22160..22375 /rpt_family="B3"
22441..22574 /rpt_family="MTC"
23116..23299 /rpt_family="B2_Mm2"
24873..24974 /rpt_family="RSINE1"

Query Match      14.0%; Score 42; DB 6; Length 223110;
Best Local Similarity 57.7%; Pred No. 0.012;
Matches 75; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 24 GTGCTACGATGAAGACTTGAAATATTTCTTCATTTGAATAAAGACGCTAACTCCCAATT 83
Db 52830 GTGCTTTTAAATATCTAATGAATTTCTTAAGTTATATAAAGCATGGTTTTTTTATT 52889

Qy 84 GTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAATAAGA 143
Db 52890 ATAGTTAAACATGTTAATTTTAAAGGAGTAATTTCTTAAAGGGCTAGTTTACTCTGAAGT 52949

Qy 144 AGTAACCCAGA 153
Db 52950 AGTAGTAAGA 52959

RESULT 15
AC131058/c
LOCUS
DEFINITION Mus musculus clone RP24-87I11, WORKING DRAFT SEQUENCE, 9 unordered
               pieces.
ACCESSION AC131058
VERSION AC131058.3 GI:28913189
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
               Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 224009)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP24-87I11
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 224009)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
               Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
               Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A.,
               Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
               Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
               Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
               Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,
               Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
               Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
               McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,

```


Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (16-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 3 (bases 1 to 224009)
AUTHORS Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Jones, C., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Levine, R., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (11-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On Mar 11, 2003 this sequence version replaced gi:28273459. All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIRB
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information
 Center project name: L27067
 Center clone name: 87_I_11

----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 22173 bases at least Q40
 Consensus quality: 222850 bases at least Q30
 Consensus quality: 223093 bases at least Q20
 Insert size: 223209; sum-of-coverage
 Quality coverage: 7.2 in Q20 bases; sum-of-coverage

* NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 10515: contig of 10515 bp in length
 * 10516 10615: gap of 100 bp
 * 10616 11771: contig of 1156 bp in length
 * 11772 11871: gap of 100 bp
 * 11872 12907: contig of 1036 bp in length
 * 12908 13007: gap of 100 bp
 * 13008 140645: contig of 127638 bp in length
 * 140646 140745: gap of 100 bp

* 140746 142799: contig of 2054 bp in length
 * 142800 142899: gap of 100 bp
 * 142900 143208: contig of 5309 bp in length
 * 148209 148308: gap of 100 bp
 * 148309 180237: contig of 31929 bp in length
 * 180238 180337: gap of 100 bp
 * 180338 222437: contig of 42100 bp in length
 * 222438 222537: gap of 100 bp
 * 222538 224009: contig of 14772 bp in length.
 * 222539 Location/Qualifiers

FEATURES
 source

1. 224009
 /organism="Mus musculus"
 /mol_type="Genomic DNA"
 /db_xref="taxon:10090"
 /clones="RP24-87111"
 /clone_lib="RPCI-24 Male Mouse BAC"

misc_feature

1. 10515
 /note="assembly_fragment"

clone_end:SP6
 vector_side:left

gap

10516..10615
 /estimated_length=100

misc_feature

10616..11771
 /note="assembly_fragment"

gap

11772..11871
 /estimated_length=100

misc_feature

11872..12907
 /note="assembly_fragment"

gap

12908..13007
 /estimated_length=100

misc_feature

13008..140645
 /note="assembly_fragment"

gap

140646..140745
 /estimated_length=100

misc_feature

140746..142799
 /note="assembly_fragment"

gap

142800..142899
 /estimated_length=100

misc_feature

142900..148208
 /note="assembly_fragment"

gap

148209..148308
 /estimated_length=100

misc_feature

148309..180237
 /note="assembly_fragment"

gap

180238..180337
 /estimated_length=100

misc_feature

180338..222437
 /note="assembly_fragment"

gap

222438..222537
 /estimated_length=100

misc_feature

222538..224009
 /note="assembly_fragment"

clone_end:T7
 vector_side:right

ORIGIN

Query Match 14.0%; Score 42; DB 12; Length 224009;
 Best Local Similarity 57.7%; Pred. No. 0.012;
 Matches 75; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 24 GTGACTACGATGAAGACTTGAATATTTCTTCATTGGAATAAACAGCTAACTCCCAATT 83

Db 98320 GTGATTTTAAATTAATCTAATGAATTTCTTAAAGTTAATAAAGCAATGTTTTTTATT 98261

Qy 84 GTACCTAAGTGGGAGTTTATCATGAAGAATGTTTAAATGCTTAATTTCAATAAGA 143

Db 98260 ATAGTTTAAACATGTTAAATTTTAAAGGAGTAAATCTTAAAGGGCTAGTTTACTCTGAAGT 98201

Qy 144 AGTAACACCA 153

Db 98200 AGTAGTAAGA 98191

```

RESULT 16
AC158640/c
LOCUS
DEFINITION
Mus musculus clone RP23-46715, WORKING DRAFT SEQUENCE, 3 unordered
pieces.
AC158640
AC158640.5 GI:63094000
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.
MUS musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 168914)
Muzny, D., Adams, C., Agbai II, O., Allen, C., Alsebrook, S., Archer, P.,
Aredondo, H., Bandaranaike, D., Banguira, L., Beltran, B., Beltran, R.,
Beraducci, A., Biswalo, K., Blyth, P., Bonham, H., Buhay, C., Burch, P.,
Cadoree, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I.,
Chacko, J., Chahrour, M., Chavez, D., Chen, A., Chen, R.,
Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A.,
Curry, S., Dai, W., Davila, M.L., Davis, C., Davy-Carroll, L., De
Anda, C., Delgado, O., Denson, S., Detamo, C., Ding, Y., Dinh, H.,
Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K.,
Druida, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M.,
Fernandez, S., Fernandez, P., Flagg, N., Forbes, L., Foster, P.,
Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T.,
Gaskin, C., Gensch, S., Ghose, S., Gill, R., Gonzalez, D.,
Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haeblerlen, K.,
Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P.,
Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J.,
Hines, S., Hitchens, M., Hodgson, A., Hogue, M., Hollins, B.,
Howell, L.T., Hui, S., Hume, J., Imo, K., Jackson, A., Jackson, L.,
Jacob, L.T., Jiang, H., Johnson, B., Johnson, R., Kalafus, K., Kelly, S.,
Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lara, F.,
Leal, S., Lee, K., Lee, S., LeGall, F.I., Lemon, S., Lewis, L., Li, B.,
Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D.,
London, P., Lopez, J., Lorensuhow, L., Lozada, R., Luk, T., Madu, R.,
Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E.,
McClelland, H., McPherson, J., Mercadao, C., Metzker, M.,
Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M.,
Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Eastaugh, E.,
Nott, A., Nwaokelumen, O., Parker, D., Pasternak, S., Patel, B.,
Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B.,
Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T.,
Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quiroz, J.,
Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S.,
Rives, C., Rodriguez, F., Rojas, A., Ruiz, S.J., Sana, M., Sanders, W.,
Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, H., Shen, Y.,
Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R.,
Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R.,
Trejos, Z., Usmani, K., Vargo, C., Verduzco, D., Villasana, D., Virk, D.,
Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J.,
Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R.,
Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L.,
Zhang, Z., Zhou, J., Weinstein, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 168914)
Worley, K.C.
Direct Submission
Submitted (23-MAR-2005) Human Genome Sequencing Center, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 168914)
Worley, K.C.
Direct Submission
Submitted (07-MAY-2005) Human Genome Sequencing Center, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 7, 2005 this sequence version replaced gi:63055078.
----- Genome Center
Center: Baylor College of Medicine
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help.tmc.edu
-----

```

```

----- Project Information
Center project name: MCGO
Center clone name: RP23-46715
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 168381 bases at least Q40
Consensus quality: 168573 bases at least Q30
Consensus quality: 168629 bases at least Q20
Estimated insert size: 173391; sum-of-contigs estimation
Estimated insert size: 168735; agarose-fp estimation
Quality coverage: 8x in Q20 bases; agarose-fp estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 22912: contig of 22912 bp in length
* 22913 23012: gap of unknown length
* 23013 63235: contig of 40223 bp in length
* 63236 63335: gap of unknown length
* 63336 168914: contig of 105579 bp in length.
FEATURES             Location/Qualifiers
     source            1..168914
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="RP23-46715"
     gap              22913..23012
                     /estimated_length=unknown
                     63236..63335
                     /estimated_length=unknown
ORIGIN
Query Match          13.9%; Score 41.8; DB 12; Length 168914;
Best Local Similarity 58.4%; Pred. No. 0.014;
Matches             0; Mismatches 52; Indels 0; Gaps 0;
QY 27 ACTACGATGAAGACTTGGAAATATTCTTCATTGAAATAAAGCAGCTAACTCCCAATTGTA 86
DB 80739 AGTACTATGAAGAGGGAATATGTGTCATTAAGAAAGTATATATTTCCAGATTATA 80680
QY 87 CCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAATTAAGAAGT 146
DB 80679 CTGAACCTCCAGGATGCTGATATAAATAATAGAGAGTTTCTCTGAATGTTTAACTTAGGAAT 80620
QY 147 AACCA 151
DB 80619 GAACA 80615
-----
RESULT 17
AC111046
LOCUS
DEFINITION
Mus musculus chromosome 10, clone RP23-479A12, complete sequence.
AC111046
AC111046.28 GI:46849819
KEYWORDS
HTG.
MUS musculus (house mouse)
SOURCE
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 231134)
Birren, B., Nusbaum, C. and Lander, E.
AUTHORS

```

| | | | |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| TITLE JOURNAL REFERENCE AUTHORS | Mus musculus chromosome 10, clone RP23-479A12 | TITLE JOURNAL COMMENT | Zimmer, A. and Zody, M. Direct Submission Submitted (29-APR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Apr 29, 2004 this sequence version replaced gi:45597676. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Center: Whitehead Institute/MIT Center for Genome Research Center code: WtBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@broad.mit.edu ----- Project Information Center project name: L19819 Center clone name: 479_A_12 ----- |
| | Unpublished | | |
| | 2 (bases 1 to 231134) | | |
| | Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamaras, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | | |
| Direct Submission Submitted (18-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 3 (bases 1 to 231134) | | | |
| REFERENCE AUTHORS | Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | TITLE JOURNAL COMMENT | Zimmer, A. and Zody, M. Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Apr 29, 2004 this sequence version replaced gi:45597676. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Center: Whitehead Institute/MIT Center for Genome Research Center code: WtBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@broad.mit.edu ----- Project Information Center project name: L19819 Center clone name: 479_A_12 ----- |
| | Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 231134) | | |
| | Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, C., Nguyen, J., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | | |
| | Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 231134) | | |
| REFERENCE AUTHORS | Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, J., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | TITLE JOURNAL COMMENT | Zimmer, A. and Zody, M. Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Apr 29, 2004 this sequence version replaced gi:45597676. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Center: Whitehead Institute/MIT Center for Genome Research Center code: WtBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@broad.mit.edu ----- Project Information Center project name: L19819 Center clone name: 479_A_12 ----- |
| | Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 231134) | | |
| | Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, J., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | | |
| | Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 231134) | | |
| REFERENCE AUTHORS | Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, J., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | TITLE JOURNAL COMMENT | Zimmer, A. and Zody, M. Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Apr 29, 2004 this sequence version replaced gi:45597676. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Center: Whitehead Institute/MIT Center for Genome Research Center code: WtBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@broad.mit.edu ----- Project Information Center project name: L19819 Center clone name: 479_A_12 ----- |
| | Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 231134) | | |
| | Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, J., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | | |
| | Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 231134) | | |
| REFERENCE AUTHORS | Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, J., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | TITLE JOURNAL COMMENT | Zimmer, A. and Zody, M. Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Apr 29, 2004 this sequence version replaced gi:45597676. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Center: Whitehead Institute/MIT Center for Genome Research Center code: WtBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@broad.mit.edu ----- Project Information Center project name: L19819 Center clone name: 479_A_12 ----- |
| | Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 231134) | | |
| | Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, J., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | | |
| | Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 231134) | | |
| REFERENCE AUTHORS | Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, J., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | TITLE JOURNAL COMMENT | Zimmer, A. and Zody, M. Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Apr 29, 2004 this sequence version replaced gi:45597676. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Center: Whitehead Institute/MIT Center for Genome Research Center code: WtBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@broad.mit.edu ----- Project Information Center project name: L19819 Center clone name: 479_A_12 ----- |
| | Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 231134) | | |
| | Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, J., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | | |
| | Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 231134) | | |
| REFERENCE AUTHORS | Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, J., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | TITLE JOURNAL COMMENT | Zimmer, A. and Zody, M. Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Apr 29, 2004 this sequence version replaced gi:45597676. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Center: Whitehead Institute/MIT Center for Genome Research Center code: WtBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@broad.mit.edu ----- Project Information Center project name: L19819 Center clone name: 479_A_12 ----- |
| | Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 231134) | | |
| | Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, J., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | | |
| | Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 231134) | | |
| REFERENCE AUTHORS | Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, J., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | TITLE JOURNAL COMMENT | Zimmer, A. and Zody, M. Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Apr 29, 2004 this sequence version replaced gi:45597676. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Center: Whitehead Institute/MIT Center for Genome Research Center code: WtBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@broad.mit.edu ----- Project Information Center project name: L19819 Center clone name: 479_A_12 ----- |
| | Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 231134) | | |
| | Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, J., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | | |
| | Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 231134) | | |
| REFERENCE AUTHORS | Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, J., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. | TITLE JOURNAL COMMENT | Zimmer, A. and Zody, M. Direct Submission Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Apr 29, 2004 this sequence version replaced gi:45597676. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Center: Whitehead Institute/MIT Center for Genome Research Center code: WtBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@broad.mit.edu ----- Project Information Center project name: L19819 Center clone name: 479_A_12 ----- |
| | Direct Submission Submitted (23-MAR-200 | | |

| | | | |
|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| repeat_region | complement(12900..13931) /rpt_family="Lx5" | REFERENCE AUTHORS | 1 (bases 1 to 237432) Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Georegeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hernandez, J., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhera, L., Loulsegod, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoleme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajd, S., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Streinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A. |
| Db | 27 ACTACGATGAGCTTGAATATTTCTTCATTTGAATTAACAGCTACTCCCAATTTGTA 86 | TITLE JOURNAL REFERENCE AUTHORS | 2 (bases 1 to 237432) Worley, K. C. |
| Db | 51592 AGTACTATGAAGAGGGAATATTTGTCTATTAAAAAAGTATTAATTTCCAGATTATA 51651 | TITLE JOURNAL | Direct Submission Unpublished |
| Qy | 87 CCTAACTAGGGAGCTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAATTAAGAAGT 146 | TITLE JOURNAL REFERENCE AUTHORS | 3 (bases 1 to 237432) Submitted (29-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA |
| Qy | 147 AACCA 151 | TITLE JOURNAL | Rat Genome Sequencing Consortium. |
| Db | 51712 GAACA 51716 | REFERENCE AUTHORS JOURNAL | Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA |
| RESULT 18 | AC123182/c | COMMENT | On Nov 15, 2002 this sequence version replaced gi:23664532. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence |
| DEFINITION | Rattus norvegicus clone CH230-239L13, *** SEQUENCING IN PROGRESS | | |
| LOCUS | AC123182 237432 bp DNA linear HTG 15-NOV-2002 | | |
| ACCESSION | AC123182 | | |
| VERSION | AC123182.4 GI:25007126 | | |
| KEYWORDS | HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED. | | |
| SOURCE | Rattus norvegicus (Norway rat) | | |
| ORGANISM | Rattus norvegicus | | |
| | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus. | | |

contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GNGG
Center clone name: CH230-239L13
----- Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 196609 bases at least Q40
Consensus quality: 201630 bases at least Q30
Consensus quality: 204680 bases at least Q20
Estimated insert size: 199959; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 221906: contig of 221906 bp in length
* 221907: gap of unknown length
* 222007 233185: contig of 11179 bp in length
* 233186 233285: gap of unknown length
* 233286 234555: contig of 1270 bp in length
* 234556 234656: gap of unknown length
* 234656 234732: contig of 2777 bp in length.

FEATURES

source
1. .237432
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-239L13"
631. .1493
/note="clone boundary"
clone_end:T7
site:
end_sequence:BZ107018"
675. .1447
/note="clone boundary"
clone_end:T7
site:
end_sequence:BZ107018"
50656. .52220
/note="wgs_end_extension"
clone_end:T7"
91278. .93768
/note="wgs_end_extension"
clone_end:T7"
164639. .169332
/note="wgs_end_extension"
clone_end:T7"
172237. .173770
/note="wgs_end_extension"
clone_end:T7"
177148. .178958
/note="wgs_end_extension"
clone_end:T7"
complement(220855..221735)
/note="clone boundary"
clone_end:Sp6
site:
end_sequence:BZ107021"

gap
gap
gap
ORIGIN

Query Match 13.9%; Score 41.6; DB 12; Length 237432;
Best Local Similarity 53.8%; Pred. No. 0.016; Mismatches 0; Gaps 0;
Matches 86; Conservative 0; Indels 74; Indels 0; Gaps 0;
QY 37 AGACCTGAATATCTTCATTGAAATAAAGACGCTAACTCCCAATTTGCTACTAGG 96
Db 73971 ATATTTAATAAATCTCTTTTGTCTTAAACCCCTCTGCTAGCACTGTGATTATAAAGT 73912
QY 97 GGAGTTTATCATGAGAAATGCTTTTAAATGCTTATTTTCAATAAGAGTAAACAGAGTG 156
Db 73911 CAAGTTTAGGACACTGTAATCTCTCAGTACAAAATATCCAAATTATTAGTTTACTTTATG 73852
QY 157 TTGCAACATGCTGTTTAAATAACCCGACAAACTTCAATCAC 196
Db 73851 TTGCTATGCTCCATATACAAACAACAACAACTTAAAAAAC 73812

RESULT 19
LOCUS CR792426/c

DEFINITION Danio rerio chromosome 3 clone DREY-23A11, WORKING DRAFT SEQUENCE, 3 unordered pieces.
ACCESSION CR792426
VERSION CR792426.13 GI:86439266

KEYWORDS HTG; HTGS PHASE1; HTGS ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 242992)
AUTHORS Smith, M.
DIRECT SUBMISSION
SUBMITTED (31-JAN-2006) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk
CONTACT: zfish-help@sanger.ac.uk
PROJECT INFORMATION
http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#dataeight
On Feb 2, 2006 this sequence version replaced gi:85814118.

COMMENT

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK23A11
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 242229 bases at least Q40
Consensus quality: 242331 bases at least Q30
Consensus quality: 242425 bases at least Q20
Insert size: 242792; sum-of-contigs
Insert size: 235334; 12.6% error; agarose-fp
Quality coverage: 7.24x in Q20 bases; sum-of-contigs Quality coverage: 7.83x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 112730: contig of 112730 bp in length

```

* 112731 112830: gap of 100 bp
* 112831 240283: contig of 127453 bp in length
* 240284 240383: gap of 100 bp
* 240384 242992: contig of 2609 bp in length.
FEATURES
    source
        1..242992
            /organism="Danio rerio"
            /mol_type="genomic DNA"
            /db_xref="RZPD:HUKGB735A11230"
            /db_xref="taxon:7955"
            /chromosome="3"
            /clone="DKEY-23A11"
            /clone_lbb="DanioKey"
        1..112730
            /note="assembly_fragment:00647"
            /fragment_chain:1
        112831..240283
            /note="assembly_fragment:02028"
            /fragment_chain:1
        240384..242992
            /note="assembly_fragment:03120.0"
ORIGIN
Query Match      13.98; Score 41.6; DB 12; Length 242992;
Best Local Similarity 50.04; Pred. No. 0.016;
Matches 104; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
QY 42 TGAATATTCCTTCATTGGAATAAACAGCTAACCTCCCAATTTGACCTAACTAGGGGAGT 101
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148000 TGAAGTGTTCACAGTATTGCAATGTGTAATGAACCTGTATAGTCTTCATTGTGCAAT 147941
QY 102 TTATCATCGAAGAAATGTTTAAATGCTTAAATTTTCAATAGAAAGTAAACAGAGTGTTCGA 161
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
147940 TTAGTACATACAATATCTTTTAAGCTACAATTTATAATTTCTTATACCAAAATAGTTCCAC 147881
QY 162 ACATGCTGTTAAATAACCCGCAACCTCAATCACTATAGCTAGTAGTGAATCTGTCATCTG 221
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
147880 AATTGCTTGAATATTATCACACAGCCTTAAAGCAGCATGCAATTTGATTAACGTTCCATTTG 147821
QY 222 CAAGGATCCCAAGTAAACCACTATTG 249
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
147820 CAATGATCCCATATTATGAACTGTG 147793

RESULT 20
AC094947/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-6C17, WORKING DRAFT SEQUENCE, 3
ACCESSION AC094947
VERSION AC094947.6 GI:30467040
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
    Sciurognathi; Muridae; Muridae; Murinae; Rattus.
REFERENCE
    1 (bases 1 to 222752)
    Muzny D, Marie, Metzker M, Lee, Abramson S, Adams C, Alder J,
    Allen C, Allen H, Albrooks S, Amin A, Anguiano D,
    Anyalebechi V, Aoyagi A, Ayodeji M, Baca E, Baden H,
    Baldwin D, Bandaranaike D, Barber M, Barnstead M, Benahmed F,
    Biswal O, Blair J, Blankenburg K, Blyth P, Brown M,
    Bryant N, Buhay C, Burch P, Burrell K, Calderon E,
    Cardenas V, Carter K, Cavazos I, Ceasar H, Chen A,
    Chacko J, Chavez D, Chen G, Chen R, Chen X, Chen Z, Chu J,
    Cleveland C, Cockrell R, Cox C, Coyle M, Cree A, D'Souza L,
    Davila M, Davis C, Davy-Carroll L, De Anda C, Dederich D,
    Delgado O, Denson S, Deramo C, Ding Y, Dinh H, Divya K,
    Draper H, Dugan-Rocha S, Dunn A, Durbin K, Duval B, Eaves K,
    Egan A, Escotto M, Eugene C, Evans C, Falls T, Fan G,
    Fernandez S, Finley M, Flagg N, Forbes L, Foster M, Foster P,
    Fraser C, Gabisi A, Ganta R, Garcia A, Garner T, Garza M,

```

Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, W. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartabeyn, A., Sisson, I., Sitter, C. D., Smajic, D., Sneed, J., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Vallas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission
 Unpublished
 2 (bases 1 to 222752)
 Worley, K. C.
 Direct Submission
 Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 222752)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On May 9, 2003 this sequence version replaced gi:24941592.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GSUG
 Center clone name: CH230-6C17
 ----- Summary Statistics
 Assembly program: Atlas;
 Consensus quality: 213463 bases at least Q40
 Consensus quality: 215000 bases at least Q30

```
Consensus quality: 216043 bases at least Q20
Estimated insert size: 223826; sum-of-Contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-Contigs estimation

*****
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: this is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*****

1 219872: contig of 219872 bp in length
* 219873 219972: gap of unknown length
* 219973 221181: contig of 1209 bp in length
* 221182 221281: gap of unknown length
* 221282 222752: contig of 1471 bp in length.
*****
Location/Qualifiers
1. .222752
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-6C17"
219873..219972
/estimated_length=unknown
221182..221281
/estimated_length=unknown

FEATURES
source
Query Match 13.8%; Score 41.4; DB 12; Length 222752;
Best Local Similarity 48.1%; Pred. No. 0.019;
Matches 117; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

Q1 58 TGAATAACAGCTAACTCCCAATTTGACCTACTAGGAGGAGTTTATCATGAGAATG 117
D1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D2 16821 TGTAAATACAGATATGTTGCAAAATATATATGTAATTTGTGCATATGCTGACAGATA 16762
Q1 118 TTTAAATGCTTAATTTCAATAAGAGTACACAGAGTGTGCAACATGCTGTTAAATAA 177
D1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D2 16761 AATATATCAATTTATGTTATATATATATATATATATATATATATATATATATA 16702
Q1 178 CCCGACAAATCTCAATCACTAGCTGAGTAGAGTGCATTCGCAAGGATCCAGAGTA 237
D1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D2 16701 CAGATTTCTATAACCTATCTGATCTGTTGCTATTTTAAATTTAAAGAACTACCTAATA 16642
Q1 238 ACCAGTATTTGGAATGCAATGTTGAACCGACCATCTAATATATCTGCTGATTAGGAAG 297
D1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D2 16641 ATTGAAGAATTATAGATATATATATATTTTACTTTTGTGATCTGTTTGTAGGTAG 16582
Q1 298 TAT 300
D1 |||||
D2 16581 GAT.16579

RESULT 21
AC131839/c
LOCUS AC131839 253693 bp DNA linear HTG 13-NOV-2002
DEFINITION Rattus norvegicus clone CH230-213, *** SEQUENCING IN PROGRESS ***,
2 unordered pieces.
ACCESSION AC131839
VERSION AC131839.3 GI:24941378
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 253693)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.B., Amarantunge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Caron,T.P., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Delaney,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Dellaway,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louissege,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sison,I.,
Sodergren,E., Sonaite,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Umani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,O.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 253693)
Worley,K.C.
Direct Submission
Submitted (26-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 253693)
Worley,K.C.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 13, 2002 this sequence version replaced gi:23664809.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: TUQL
Center clone name: CH230-213
----- Summary Statistics
```


in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: TUQU
 Center clone name: CH230-2C7
 ----- Summary Statistics

Assembly program: Phrap; version 0.990329
 Consensus quality: 254993 bases at least Q40
 Consensus quality: 259665 bases at least Q30
 Consensus quality: 262708 bases at least Q20
 Estimated insert size: 273866; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
 * NOTE: This sequence may represent more than one clone.
 * NOTE: This is a 'working draft' sequence. It currently consists of 11 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 5545: contig of 5545 bp in length
 * 5546 5645: gap of unknown length
 * 5646 10296: contig of 4651 bp in length
 * 10297 10396: gap of unknown length
 * 10397 258638: contig of 248242 bp in length
 * 258639 258738: gap of unknown length
 * 258739 260235: contig of 1497 bp in length
 * 260236 260335: gap of unknown length
 * 260336 261450: contig of 1115 bp in length
 * 261451 261550: gap of unknown length
 * 261551 262936: contig of 1386 bp in length
 * 262937 263036: gap of unknown length
 * 263037 263499: contig of 1313 bp in length
 * 264450 264459: gap of unknown length
 * 264450 265703: contig of 1254 bp in length
 * 265704 265803: gap of unknown length
 * 265804 267592: contig of 1789 bp in length
 * 267593 267692: gap of unknown length
 * 267693 269812: contig of 2120 bp in length
 * 269813 271927: gap of unknown length
 * 269913 271927: contig of 2015 bp in length.

FEATURES
 source
 1. 271927
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-2C7"
 5546..5645
 /estimated_length=unknown
 8661..10296
 /note="wgs contig"
 10297..10396
 /estimated_length=unknown
 258639..258738
 /estimated_length=unknown
 260236..260335
 /estimated_length=unknown

gap
 misc_feature
 gap
 gap
 gap

gap
 gap
 gap
 gap
 gap
 gap
 gap
 ORIGIN

Query Match 13.8%; Score 41.4; DB 12; Length 271927;
 Best Local Similarity 48.1%; Pred. No. 0.019;
 Matches 117; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
 QY 58 TGAATAAACAGCTAACTCCCAAAATTGACCTAAGTGGGAGTTTATCATGAAGAAATG 117
 Db 49459 TGTAAACAGATATTGTTGCAAAATTTATATGTAATTTGTCATATGCTGCACAGATA 49518
 QY 118 TTTAAATGCTTAATTTTCAATAAGACTAACCCAGAGTTTGCAACATGCTGTTAAATAA 177
 Db 49519 AATATATCATTAATGTTATATATATATATATATATATATATATATATATATATAA 49578
 QY 178 CCCGACAACTTCAATCACTATAGCTGTAGTAGTGCATTTCTCAAGGATCCCAAGATA 237
 Db 49579 CAGATTTCTATAACCTATCTGATCTGTTGCTATTTTAAATTAAGAACTACCTAATA 49638
 QY 238 ACCAGTATTTTGGAAATGCAATGTTGAACCCAGCACTACTAATATCTCTGATTAAGGAAG 297
 Db 49639 ATTGAAGAAATTAGATATTATTAATTTTACCTTTTGTGTTTGTGTTTGTAGTAGTAG 49698
 QY 298 TAT 300
 Db 49699 GAT 49701

RESULT 23
 LOCUS AP004980/c 23709 bp DNA linear PLN 22-JUL-2003
 DEFINITION Lotus corniculatus var. japonicus genomic DNA, chromosome 3,
 clone:Lfj13N17, TM0436a, complete sequence.
 ACCESSION AP004980
 VERSION AP004980.1 GI:21907998
 KEYWORDS HTG.
 SOURCE Lotus corniculatus var. japonicus (Lotus japonicus)
 ORGANISM Lotus corniculatus var. japonicus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
 Lotus.
 REFERENCE 1
 AUTHORS Kaneko,T., Nakamura,Y., Asamizu,E., Kato,T., Sato,S. and Tabata,S.
 TITLE Structural Analysis of a Lotus japonicus Genome. I. Sequence
 Features and Mapping of Sixty-six TAC clones which cover the 6.7 Mb
 Regions of the Genome
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 23709)
 AUTHORS Nakamura,Y.
 TITLE Direct Submission
 JOURNAL Submitted (26-MAR-2002) Yasukazu Nakamura, Kazusa DNA Research
 Institute, Department of Plant Gene Research; 1532-3, Yana,
 Kisarazu, Chiba 292-0812, Japan (E-mail:yn@kzusa.or.jp,
 URL:<http://www.kazusa.or.jp>, Tel:81-438-52-3935,
 Fax:81-438-52-3934)
 FEATURES
 source
 1. 23709
 Location/Qualifiers
 /organism="Lotus corniculatus var. japonicus"
 /mol_type="genomic DNA"
 /variety="japonicus"
 /db_xref="taxon:34305"

```

/chrnosome="3"
/clone="Lj113N17"
/note="TM0436a, a part of TAC clone:TM0436.
synonym: Lotus japonicus"

Query Match      13.6%; Score 40.8; DB 4; Length 23709;
Best Local Similarity 49.5%; Pred. No. 0.031;
Matches 105; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

ORIGIN
1
4 ATTGACAGGCTAAATGCTAGCTACGATGAGACTTGAATATTTCTTCATTGGAAT 63
10080 ATTGGTTATCTAAATGAGAATTTCAAAAAAAGAGCAATACCTCATCTAAATCTGTCT 10021
64 AAACAGCTAACTCCCAATTTGTACTAACTAGGAGCTTTATCATGAAGAAATGTTTAAA 123
10020 AAAATATTTTCCCATTTTACAAAATTTAGGGAATAATAAAAAATATAAAATCTAT 9961
124 TGCTTAATTTTCAATTAAGATTAACGAGATGTTGCAACATGCTGTTAAATAACCCGAC 183
9960 TATATAAATTAAGAGAAGACTCAACAAAAAAGTGCAATGTTATCTTCCATTCATGAC 9901
184 AACTTCAATCACTATAGCTGTAGTAGTGC 215
9900 AATGCTCATGTTGCTGTGTGTGTGAATGC 9869

RESULT 24
LOCUS      BX255970      158755 bp      DNA      linear      VRT 17-APR-2005
DEFINITION Zebrafish DNA sequence from clone CH211-136G2 in linkage group 4
            Contains the gene for a novel slow skeletal troponin T family
            protein, the gene for a novel protein (zgc:85809), seventeen novel
            genes and a CpG island, complete sequence.
ACCESSION  BX255970
VERSION     BX255970.8  GI:49203508
KEYWORDS   HTG.
SOURCE      Danio rerio (zebrafish)
            Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 158755)
            Direct Submission
            Submitted (02-DEC-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests:
            http://www.sanger.ac.uk/Projects/D_rerio/faqs.shtml#dataeight
            On Jun 24, 2004 this sequence version replaced gi:38228952.
            ----- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -----
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            CH211-136G2 is from a CHORI-211 BAC library
            VECTOR: pPARBAC2.1
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
Any regions longer than 1kb tagged as misc-feature 'unsure' are
part of a tandem repeat of more than 10kb in length where it has
not been possible to anchor the base differences between repeat
copies. The region has been built up based on the repeat element
to match the total size of repeat indicated by restriction digest,
but repeat copies may not be in the correct order and the usual
finishing criteria may not apply.
Location/Qualifiers
1. 158755
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/chrnosome="4"
/clone="CH211-136G2"
/clone_lib="CHORI-211"
1
/note="Clone left end: CH211-136G2"
join(836..882,4803..4853,7092..7096,11651..11672,
11766..11823,14755..14871,15177..15254,15358..15471,
15805..15917,16987..17077,18279..18319,24154..24209)
/locus_tag="CH211-136G2.1-002"
join(836..882,4803..4853,7092..7096,11651..11672,
11766..11823,14755..14871,15177..15254,15358..15471,
15805..15917,16987..17077,18279..18319,24154..24209)
/locus_tag="CH211-136G2.1-002"
/product="novel slow skeletal troponin T family protein"
/note="match: ESTs: Em:CF347331.1 Em:CF348408.1
Em:CK142922.1"
join(840..882,4803..4853,7092..7096,11079..11105,
11651..11672,11766..11823,14755..14871,15177..15254,
15358..15471,15805..15917,16987..17077,18279..18312)
/locus_tag="CH211-136G2.1-004"
/product="novel slow skeletal troponin T family protein"
/note="match: ESTs: Em:CN504044.1"
join(848..882,4803..4853,7092..7096,11651..11672,
11766..11823,14755..14871,15177..15254,15358..15471,
15805..15917,16987..17077,18517..18554,24154..24411)
/locus_tag="CH211-136G2.1-001"
join(848..882,4803..4853,7092..7096,11651..11672,
11766..11823,14755..14871,15177..15254,15358..15471,
15805..15917,16987..17077,18517..18554,24154..24411)
/locus_tag="CH211-136G2.1-001"
/product="novel slow skeletal troponin T family protein"
/note="match: ESTs: Em:CF66249.1 Em:CK029104.1"
join(4806..4853,7092..7096,8006..8026,9552..9557,
9805..9825,9977..9997,10413..10430,10527..10550,
11079..11105,11651..11672,11766..11823,14755..14871,
15177..15254,15358..15471,15805..15917,16987..17077,
18517..18554,24154..24199)
/locus_tag="CH211-136G2.1-003"
join(4806..4853,7092..7096,8006..8026,9552..9557,
9805..9825,9977..9997,10413..10430,10527..10550,
11079..11105,11651..11672,11766..11823,14755..14871,
15177..15254,15358..15471,15805..15917,16987..17077,
18517..18554,24154..24199)
/locus_tag="CH211-136G2.1-003"
/product="novel slow skeletal troponin T family protein"
/note="match: ESTs: Em:CN508154.1"
join(4816..4853,7092..7096,8006..8026,9552..9557,
9805..9825,9977..9997,10413..10430,10527..10550,
11079..11105,11651..11672,11766..11823,14755..14871,
15177..15254,15358..15471,15805..15917,16987..17077,
18517..18554,24154..24199)
/locus_tag="CH211-136G2.1-003"
/standard_name="OTTDARF0000007299"
/note="match: proteins: Tr:O13096"
/codon_start=1
FEATURES
source
1. 158755
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/chrnosome="4"
/clone="CH211-136G2"
/clone_lib="CHORI-211"
1
misc_feature
1
/note="Clone left end: CH211-136G2"
join(836..882,4803..4853,7092..7096,11651..11672,
11766..11823,14755..14871,15177..15254,15358..15471,
15805..15917,16987..17077,18279..18319,24154..24209)
/locus_tag="CH211-136G2.1-002"
join(836..882,4803..4853,7092..7096,11651..11672,
11766..11823,14755..14871,15177..15254,15358..15471,
15805..15917,16987..17077,18279..18319,24154..24209)
/locus_tag="CH211-136G2.1-002"
/product="novel slow skeletal troponin T family protein"
/note="match: ESTs: Em:CF347331.1 Em:CF348408.1
Em:CK142922.1"
join(840..882,4803..4853,7092..7096,11079..11105,
11651..11672,11766..11823,14755..14871,15177..15254,
15358..15471,15805..15917,16987..17077,18279..18312)
/locus_tag="CH211-136G2.1-004"
/product="novel slow skeletal troponin T family protein"
/note="match: ESTs: Em:CN504044.1"
join(848..882,4803..4853,7092..7096,11651..11672,
11766..11823,14755..14871,15177..15254,15358..15471,
15805..15917,16987..17077,18517..18554,24154..24411)
/locus_tag="CH211-136G2.1-001"
join(848..882,4803..4853,7092..7096,11651..11672,
11766..11823,14755..14871,15177..15254,15358..15471,
15805..15917,16987..17077,18517..18554,24154..24411)
/locus_tag="CH211-136G2.1-001"
/product="novel slow skeletal troponin T family protein"
/note="match: ESTs: Em:CF66249.1 Em:CK029104.1"
join(4806..4853,7092..7096,8006..8026,9552..9557,
9805..9825,9977..9997,10413..10430,10527..10550,
11079..11105,11651..11672,11766..11823,14755..14871,
15177..15254,15358..15471,15805..15917,16987..17077,
18517..18554,24154..24199)
/locus_tag="CH211-136G2.1-003"
join(4806..4853,7092..7096,8006..8026,9552..9557,
9805..9825,9977..9997,10413..10430,10527..10550,
11079..11105,11651..11672,11766..11823,14755..14871,
15177..15254,15358..15471,15805..15917,16987..17077,
18517..18554,24154..24199)
/locus_tag="CH211-136G2.1-003"
/product="novel slow skeletal troponin T family protein"
/note="match: ESTs: Em:CN508154.1"
join(4816..4853,7092..7096,8006..8026,9552..9557,
9805..9825,9977..9997,10413..10430,10527..10550,
11079..11105,11651..11672,11766..11823,14755..14871,
15177..15254,15358..15471,15805..15917,16987..17077,
18517..18554,24154..24199)
/locus_tag="CH211-136G2.1-003"
/standard_name="OTTDARF0000007299"
/note="match: proteins: Tr:O13096"
/codon_start=1

```

| | | |
|-----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| mRNA | /locus_tag="CH211-136G2.3-001" | |
| | join(31521..31816,34657..34755,34931..35208) | |
| | /locus_tag="CH211-136G2.3-001" | |
| | /product="novel transcript" | |
| gene | /note="match: ESTs: Em:CN174698.1" | |
| | join(36670..36967,38896..38994,39172..39449) | |
| mRNA | /locus_tag="CH211-136G2.4-001" | |
| | join(36670..36967,38896..38994,39172..39449) | |
| | /locus_tag="CH211-136G2.4-001" | |
| | /product="novel transcript" | |
| | /note="match: ESTs: Em:CN174698.1" | |
| Query Match | 13.6%; Score 40.8; DB 11; Length 158755; | |
| Best Local Similarity | 53.0%; Pred. No. 0.031; | |
| Matches | 87; Conservative 0; Mismatches 77; Indels 0; Gaps 0; | |
| QY | 36 AAGACTTGAATATCTTCATTTCGAATAACACGTAACCTCCCAATTTGTACCTTAACTAG 95 | |
| Db | 150328 AATATTGCAATATTTCAACTTCAAAAAATAACTTTGGTCACACTTGTACATTTTAAT 150387 | |
| QY | 96 GGGAGTTTATCATCGAAGAAATGTTTAAATGCTTAAATTTCAAATAAGAAGTAACACAGCT 155 | |
| Db | 150388 TTATCATACATAAATCTAGTATTTATACCATATATTTTAATTAACACATTACTTAAG 150447 | |
| QY | 156 GTTGCAACATGCTGTTTAAATAACCCGACAACTTCAATCACTAT 199 | |
| Db | 150448 TACAATACATTTAGTTAAAAAAAATAACAAATTTAAAAACAGTAT 150491 | |
| RESULT 25 | | |
| AL356073/c | | |
| LOCUS | AL356073 165917 bp DNA linear PRI 18-MAY-2005 | |
| DEFINITION | Human DNA sequence from clone RP11-81303 on chromosome 10, complete sequence. | |
| ACCESSION | AL356073 | |
| VERSION | AL356073.15 GI:16943998 | |
| KEYWORDS | HTG. | |
| SOURCE | Homo sapiens | |
| ORGANISM | Homo sapiens (human) | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo. | |
| REFERENCE | 1 (bases 1 to 165917) | |
| AUTHORS | Pelan,S. | |
| TITLE | Direct Submission | |
| JOURNAL | Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk | |
| COMMENT | Clone requests: clonerequest@sanger.ac.uk On Nov 15, 2001 this sequence version replaced gi:15590948. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/chr10 RP11-81303 is from the library RPCI-11.3 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pBACe3.6 | |
| | ----- Genome Center | |
| | Center: Wellcome Trust Sanger Institute | |
| | Center code: SC | |
| | Web site: http://www.sanger.ac.uk | |
| | Contact: vegas@sanger.ac.uk | |
| | ----- | |
| | This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least | |

one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

```

FEATURES
    source
        1..165917
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="10"
            /clone="RP11-81303"
            /clone_lib="RPC1-11.3"
    misc_feature
        1 /note="Clone_left_end: RP11-81303"
    misc_feature
        46711..46804
            /note="Sequence from uni-directional dGTP big dye
            terminator reads only."
    misc_feature
        46711..46803
            /note="Sequence from clone PCR only."
    misc_feature
        46804
            /note="Gap size estimated to be approximately 330bp by
            restriction digest data."
    misc_feature
        51135
            /note="Clone_right_end: RP11-765C10"
    misc_feature
        66101
            /note="Clone_left_end: RP11-129G17"
    misc_feature
        165917
            /note="Clone_right_end: RP11-81303"
ORIGIN
Query Match      13.6%; Score 40.8; DB 5; Length 165917;
Best Local Similarity 55.7%; Pred. No. 0.031;
Matches 78; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 33 ATGAAGACTTGAAATATTTCTTCATTTGAATAAAGACAGTAACTCCCAATTTGTACCTAAC 92
DB 55035 ATTAGGTATAGATACTTCTTATCATAGAAGTAGTACAAATCTACATCA 54976
QY 93 TAGGGGAGTTATCATGAAGAAATGTTTAAATGCTTAATTTCAATPAAGAAGTAAACAG 152
DB 54975 AAATGCTTCTGTAATGGTGAATGTTTAAATATTTACATTTGAGATCAGAAAAAGGCA 54916
QY 153 AGTGTTCACATCTCTTTA 172
DB 54915 AGGATGAAGCTTGCTTTCA 54896

RESULT 26
AC135483/c
LOCUS AC135483 166494 bp DNA linear HTG 20-NOV-2002
DEFINITION Rattus norvegicus clone CH230-261E14, *** SEQUENCING IN PROGRESS
***, 3 unordered pieces.
ACCESSION AC135483
VERSION AC135483.2 GI:25139065
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 166494)
Muzny,D,Maric,E, Metzker,M, Lee,, Abramzon,S,, Adams,C,, Alder,J,,
Allen,C,, Allen,H,, Alsbrooks,S,, Amin,A,, Anguiano,D,,
Anyalebechi,V,, Ayoyagi,A,, Ayodeji,M,, Baca,E,, Baden,H,,
Baldwin,D,, Bandaranaike,D,, Barber,M,, Barnstead,M,, Benahmed,F,,
Blawolo,K,, Blair,J,, Blankenburg,K,, Blyth,P,, Brown,M,,
Bryant,N,, Buhay,C,, Burch,P,, Burrell,K,, Calderon,E,,
Cardenas,V,, Carter,K,, Cavazos,I,, Cesar,H,, Chen,A,,
Chacko,J,, Chavez,D,, Chen,G,, Chen,Y,, Chen,Z,, Chu,J,,
Cleveland,C,, Cockrell,R,, Cox,C,, Coyle,M,, Cree,A,, D'Souza,L,,
Davila,M,L,, Davis,C,, Davy-Carroll,L,, De Anda,C,, Dederich,D,,
Delgado,O,, Denson,S,, Deramo,C,, Ding,Y,, Dinh,H,, Divya,K,,
Draper,H,, Dugan-Rocha,S,, Dunn,A,, Durbin,K,, Duval,B,, Eaves,K,,
Egan,A,, Escotto,M,, Eugene,C,, Evans,C,A,, Falls,T,, Fan,G,,
Fernandez,S,, Finley,M,, Flagg,N,, Forbes,L,, Foster,M,, Foster,P,,

```

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebrgeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoques,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,J., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhea,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J.J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwakoelameh,O., Okwuonu,G., Olarnpunsegon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Primus,E., Pu,L.-L., Plopper,F., Polindexter,A., Popovic,D., Reeves,K., Regier,M.A., Reigh,R., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Rejos,Z., Usmani,K., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smays,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,K., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhauser,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 166494)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (17-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 166494)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:24080623.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KBWY
Center clone name: CH230-261E14
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 151587 bases at least Q40

Consensus quality: 154840 bases at least Q30
 Consensus quality: 15515 bases at least Q20
 Estimated insert size: 155596; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 145399: contig of 145399 bp in length
 * 145400: gap of unknown length
 * 145500: contig of 19262 bp in length
 * 164761: gap of unknown length
 * 164762: gap of unknown length
 * 164862: contig of 1633 bp in length.

FEATURES

source

1. 166494

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-261B14"

3174..3995

/note="clone boundary"

clone_end:Sp6

site:

end_sequence:BZ245159"

72113..72978

/note="clone boundary"

clone_end:T7

site:

end_sequence:BZ245158"

75147..76243

/note="wgs_end_extension"

clone_end:T7"

76585..80486

/note="wgs_end_extension"

clone_end:T7"

164762..164861

/estimated_length=unknown

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

AC142454/c

LOCUS

DEFINITION

AC142454

Mus musculus BAC clone RP23-313G1 from chromosome 3, complete

sequence.

AC142454

AC142454.4 GI:71533458

HTG.

Mus musculus (house mouse)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC142454

Mus musculus BAC clone RP23-313G1 from chromosome 3, complete

sequence.

AC142454

AC142454.4 GI:71533458

HTG.

Mus musculus (house mouse)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Jul 30, 2005 this sequence version replaced gi:57790245.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu

Contact: submissions@watson.wustl.edu

----- Summary Statistics

Center project name: M_BA0313G01

NOTICE:

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e. phred quality
 >=30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone, fosmid clone or direct clone walk sequence.
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to obtain the consensus sequence. The
 assembly was confirmed by restriction digest.
 This finishing standard has slightly changed from the previous
 Human standard. Specifically, standards for regions of low sequence
 complexity (such as dinucleotide repeats and small unit tandem
 repeats) have been relaxed. These regions are very prevalent in the
 mouse genome, and the return on extended finishing efforts is
 minimal.
 If a sequence meets the criteria of the above statement, it needs
 no comments or tags. If the criteria are not met, such as ambiguous
 bases, then the region is duly annotated.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
 http://genome.wustl.edu

SOURCE INFORMATION:

The BAC Library has been constructed by Kazutoyo Oseawa and
 Minako Tateno in the laboratory of Pieter de Jong
 (http://www.chori.org) from female C57BL/6J mouse kidney and/or

Query Match 13.5%; Score 40.4; DB 12; Length 166494;
 Best Local Similarity 48.7%; Pred No. 0.042;
 Matches 110; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
 22 AAGTGACTACGATGAGCTTGAATATTTCTTCATTGGAATATAACAGCTACCTCCAAA 81
 DB 122057 AACTGAGAAATATATAAAATCAAAAGTTGTGGAATGTCATAGGAACCTACAGAGGCAAG 121998
 82 TTGTACCTAACAGGAGGATTTATCATGAGAAATGTTTAAATGCTTAAATTTTCAATATA 141
 DB 121997 TTTTATCAAAAGAAATPAAGACAGATGGAAGAGAAATCTCAGGCTCAGAATATACAAAAGA 121938
 142 GAAGTAAACGAGCTGTCACATCTGTTAAATAACCGCAAACTTCAATCACTATAG 201
 DB 121937 AAAGTGATATATTTGGTCAAGAAATGTTGAAATTTGAACAAAACCTTCTAGCAGAAAAG 121878
 202 CTGTAGTAGAGTGCTTCTGCAAGAGTCCAGAGTAACCACTATTT 247
 DB 121877 ATCCAGAAATCCAAACCTATATAAAGACCAAACTTACAAATAATT 121832

brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

This sequence is the entire insert of the clone.

```

FEATURES             source
    source
        1. .183223
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /chromosome="3"
            /clone="RP23-313G1"
            /clone_lib="RPCI-23"
            /clone_id="114411"
        misc_feature
            114260..114411
            /note="Sequence derived from PCR product of project DNA"

ORIGIN
Query Match      13.5%; Score 40.4; DB 6; Length 183223;
Best Local Similarity 52.4%; Pred. No. 0.042;
Matches 89; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 60 AATAAAGACAGTCACTCCCAATGTACTAACTAGGAGGAGTTTATCATGAAGAAATGTT 119
DB 7285 AACTGAACACACTCTTCTCAATGATACCTTGGTCAAGGAGGAGTAAGAAAAATGA 7226
QY 120 TAAATGCTTAATTTCAATAAGAGTAACAGAGTGTGCAACATCTGCTTAAATTAACC 179
DB 7225 GGACTTTTGAAGTTCAATGAATAAGACACAGCATACCCAAATTTATTGACATAATG 7166
QY 180 CGACAACATCACTACTATAGCTGTAGTAGAGTGCGATTTCGCAAGATC 229
DB 7165 AAAGGATTTTAAAGAGAAACTCATAGCTGAGTGCTCCAAAAAGAAAC 7116

RESULT 28
AC139360/c
LOCUS      AC139360      186836 bp      DNA      linear      ROD 20-FEB-2004
DEFINITION Mus musculus chromosome 3, clone RP23-324F4, complete sequence.
ACCESSION AC139360
VERSION    AC139360.8 GI:42716198
KEYWORDS   HTG.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 186836)
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus chromosome 3, clone RP23-324F4
Unpublished

TITLE      Location/Qualifiers
AUTHORS    1. .186836
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /chromosome="3"
            /map="3"
            /clones="RP23-324F4"
            /clone_lib="RPCI-23 Female Mouse BAC"
            /note="clone_boundary"

misc_feature 1. .6
            /note="clone_boundary"

FEATURES
source
    source
        1. .186836
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /chromosome="3"
            /map="3"
            /clones="RP23-324F4"
            /clone_lib="RPCI-23 Female Mouse BAC"
            /note="clone_boundary"

misc_feature 1. .6
            /note="clone_boundary"

```

REFERENCE AUTHORS

Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 186836)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collimore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Menes,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-JAN-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 186836)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collimore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Menes,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-FEB-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 20, 2004 this sequence version replaced gi:41058797.
All repeats were identified using RepeatMasker:
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
Smith, A.F.A. & Green, P. (1996-1997)

TITLE JOURNAL COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www.seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L19543
Center clone name: 324_F_4
----- Location/Qualifiers
1. .186836
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="3"
/map="3"
/clones="RP23-324F4"
/clone_lib="RPCI-23 Female Mouse BAC"
1. .6
/note="clone_boundary"

JOURNAL Submitted (08-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On Feb 21, 2003 this sequence version replaced gi:27819585.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@watson.wustl.edu
 ----- Summary Statistics
 ----- Center project name: M_BA0187E03

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Oseagawa and Minako Tatenio in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC124440.

FEATURES

source

1. .192918
 location/Qualifiers
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="3"
 /map="3"

/clone="RP23-187E3"
 /clone_lib="RPCI-23"

784. .1161

/rpt_family="ERVK"

1827. .2046

/rpt_family="B4"

2408. .3581

/rpt_family="L1"

4262. .4541

/rpt_family="MaLR"

4545. .4913

/rpt_family="MaLR"

7970. .8111

/rpt_family="L2"

9273. .9466

/rpt_family="B4"

10761. .10836

/rpt_family="L1"

10809. .10940

/rpt_family="L1"

12991. .13202

/rpt_family="B2"

25930. .26260

/rpt_family="L1"

repeat_region
 /rpt_family="B4"
 29078. .29212
 repeat_region
 /rpt_family="RMER15"
 30956. .31120
 repeat_region
 /rpt_family="B2"
 32294. .32449
 repeat_region
 /rpt_family="ERV1"
 32937. .32989
 repeat_region
 /rpt_family="B4"
 33851. .33968
 repeat_region
 /rpt_family="B4"
 34965. .35371
 repeat_region
 /rpt_family="MaLR"
 35375. .36428
 repeat_region
 /rpt_family="MaLR"
 36429. .36737
 repeat_region
 /rpt_family="MaLR"
 38654. .38794
 repeat_region
 /rpt_family="Alu"
 39154. .39237
 repeat_region
 /rpt_family="Alu"
 41191. .41297
 repeat_region
 /rpt_family="Alu"
 51184. .51569
 repeat_region
 /rpt_family="MaLR"
 51730. .51952
 repeat_region
 /rpt_family="L1"
 52503. .52672
 repeat_region
 /rpt_family="B4"
 52775. .52872
 repeat_region
 /rpt_family="MIR"
 53927. .54116
 repeat_region
 /rpt_family="B2"
 54118. .54230
 repeat_region
 /rpt_family="B4"
 63831. .64645
 repeat_region
 /rpt_family="RMER1B"
 64676. .64821
 repeat_region
 /rpt_family="L1"
 64899. .64946
 repeat_region
 /rpt_family="ID"
 65886. .66372
 repeat_region
 /rpt_family="ERVK"
 66375. .66543
 repeat_region
 /rpt_family="ERVK"
 66549. .66642
 repeat_region
 /rpt_family="ERVK"
 66861. .66968
 repeat_region
 /rpt_family="ERVK"
 67443. .67522
 repeat_region
 /rpt_family="ERVK"
 67853. .67913
 repeat_region
 /rpt_family="ERVK"
 67952. .68163
 repeat_region
 /rpt_family="ERVK"
 68202. .68244
 repeat_region
 /rpt_family="ERVK"
 68247. .68429
 repeat_region
 /rpt_family="ERVK"
 68432. .68918
 repeat_region
 /rpt_family="ERVK"
 68921. .69089
 repeat_region
 /rpt_family="ERVK"
 69095. .69188
 repeat_region
 /rpt_family="ERVK"
 69193. .69270
 repeat_region
 /rpt_family="L1"
 69722. .69861
 repeat_region
 /rpt_family="B4"
 72101. .72229
 repeat_region
 /rpt_family="ERV1"
 72279. .72534
 repeat_region
 /rpt_family="ERV1"
 72540. .72588


```

/rpt_family="ERV1"
72698. 72779
/rpt_family="ERV1"
72916. 72964
/rpt_family="ERV1"
74972. 75180
/rpt_family="B2"
75380. 75503
/rpt_family="L2"
75646. 75780
/rpt_family="B4"
76242. 76457
/rpt_family="B2"
78877. 79112
/rpt_family="L1"
79561. 79651
/rpt_family="L1"
79952. 80122

Query Match 13.5% Score 40.4; DB 6; Length 192918;
Best Local Similarity 52.4%; Pred. No. 0.042;
Matches 89; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 60 AAATAAAGCTTAACCTCCCAATTTCTACTACTAGGAGTTTATCATGAGAAATGTT 119
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
17872 AACTGAACAACACTCTTCTCAATGATACCTTGGTCAAGGAAGGAATAAGAAATAATGA 17813

QY 120 TAATGCTTAATTTTCAATAGAGTAACACAGAGTGTGCAACATGCTGTTAAATAACC 179
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
17812 GGACTTTTGAAGTTCATGAAATGAAGCCACGACATACCAATTTATTTGACATAATG 17753

QY 180 CGACAACTTCAATCACTATAGCTGTAGTAGAGTGCATTTCTGCAAGGATC 229
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
17752 AAAGGATTTTAAAGAGGAAACTCATAGCTGAGTGCCTCCAAAGAAAC 17703

RESULT 30
AC123390
LOCUS
DEFINITION Rattus norvegicus clone CH230-514P9, *** SEQUENCING IN PROGRESS
***, 4 unordered pieces.
AC123390
VERSION AC123390.4 GI:25137939
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Rattus.
1 (bases 1 to 213458)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.B., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,I., Garza,M.,
Gebregregis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,B., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,

```

```

Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poidexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umami,K.,
Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 213458)
Worley,K.C.
Direct Submission
Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 213458)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:22856263.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center of Medicine
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GWQS
Center clone name: CH230-514P9
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 187478 bases at least Q40
Consensus quality: 190243 bases at least Q30
Consensus quality: 191914 bases at least Q20
Estimated insert size: 192725; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces

```

* is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will be preserved

* 1 208908: contig of 208908 bp in length
 * 208909 209008: gap of unknown length
 * 209009 210058: contig of 1050 bp in length
 * 210059 210158: gap of unknown length
 * 210159 211674: contig of 1516 bp in length
 * 211675 211774: gap of unknown length
 * 211775 213458: contig of 1684 bp in length.

FEATURES

source
 1. 213458
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clones="CH230-514P9"
 misc_feature
 1. 1230
 /note="wgs contig"
 misc_feature
 2053. 3154
 /note="wgs contig"
 gap
 208909..209008
 /estimated_length=unknown
 gap
 210059..210158
 /estimated_length=unknown
 gap
 211675..211774
 /estimated_length=unknown

ORIGIN

Query Match 13.5%; Score 40.4; DB 12; Length 213458;
 Best Local Similarity 48.7%; Pred. No. 0.041;
 Matches 110; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
 QY 22 AAGTGACTGACGACGACTGTAATATCTTCATTTGTAATACACGCTAACCTCCCAA 81
 Db 126364 AACTGGAATATATATAAAATCAAAATGTGGAAATGTCAATAGGAAGCTCAGAGGCAAG 126423
 QY 82 TTGTACCTAACTAGGGGAGTTTATCATGAAGAAATGTTTAAATCCTTAAATTTTCAAAATAA 141
 Db 126424 TTTTATCAAGAATACAGACAGATGGAAGAAATCTCAGGCTCAGATATACAAAGA 126483
 QY 142 GAAGTACACGAGTGTTGCAACATGCTGTTAAATACCCGCAAACTTCAATCTATGAT 201
 Db 126484 AAAGTGAATATATTGGTCAAGAAATGTTGAAATTTGAACAAAACTTCTAGCACAAAAG 126543
 QY 202 CTGTAAGTAGTGATCTTCTCAAGGATCCCAAGAGTAACCACTATTT 247
 Db 126544 ATCCAGAAATCCAAACCTATATAAAGACCAACCTCAATAATTT 126589

RESULT 31

AC095434/c
 LOCUS AC095434 222392 bp DNA linear HTG 09-MAY-2003
 DEFINITION Rattus norvegicus clone CH230-6P20, *** SEQUENCING IN PROGRESS ***.
 ACCESSION AC095434
 VERSION AC095434.9 GI:30467849
 KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE

1 (bases 1 to 222392)
 Muzny D, Marie, Metzker M, Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Cantu, R., Garcia, A., Garner, T., Garza, M., Georegeorgis, E., Geer, K., Gill, R., Grady, W., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hoques, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lohensuhewa, L., Loulsegad, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavijevic, A., Miner, G., Minja, E., Montemayor, J., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajd, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, D., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission

Unpublished

2 (bases 1 to 222392)

Worley, K.C.

Direct Submission

Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 222392)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:24941161.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

```

----- Project Information
Center project name: GCGI
Center clone name: CH230-6P20
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 189423 bases at least Q40
Consensus quality: 192508 bases at least Q30
Consensus quality: 194817 bases at least Q20
Estimated insert size: 199567; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 1 contigs. Gaps between the contigs
  * are represented as runs of N. The order of the pieces
  * is believed to be correct as given, however the sizes
  * of the gaps between them are based on estimates that have
  * provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 22392: contig of 22392 bp in length.
FEATURES             Location/Qualifiers
     source            1..22392
                        /organism="Rattus norvegicus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10116"
                        /clone="CH230-6P20"
     misc_feature      1..1189
                        /note="wgs_contig"
     misc_feature      211317..213363
                        /note="wgs_contig"
ORIGIN
Query Match          13.5%; Score 40.4; DB 12; Length 22392;
Best Local Similarity 47.9%; Pred. No. 0.041;
Matches 116; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
59 GAAATAAAGCAGTAACTCCCAATGTTACCTAACTAGGGAGTTTATCATGAAGAATGT 118
23487 GTAATACAGATATGTTGCAAAATTTATGTAATTTGCCATATGCTGCACAGATAA 23428
119 TTAATGCTTAATTTTCAAAATGAAGTAACAGAGTGTTCACATCTGTTAAATAAC 178
23427 ATATATCATTTAATGTTTATAATATATACATACATGATATCATATATATAAC 23368
179 CCGACAACTTCAATCATATAGCTGTAGTAGTGCAATTCGCAAGATCCAGAGTAA 238
23367 AGATTTCTATAACCTATCTGATCTGTTGCTATTATTTAAAGAACTACCTAATA 23308
239 CCAGTATTTTGGAAATGCAATGTTGCAACCGACCATTAATTTATCTGCTGATTAGGAAGT 298
23307 TTGAAGATTTATAGTATTAATTTTACTTTTAGTTGTTGTTGATTAGGTAGG 23248
299 AT 300
23247 AT 23246
23247 AT 23246
RESULT 32
AE015927_05
WPCOMMENT
Sequence split into 28 fragments LOCUS AE015927 Accession AE015927
Fragment Name      Begin      End
AE015927_00        1         110000
AE015927_01       100001     210000
AE015927_02       200001     310000
AE015927_03       300001     410000
AE015927_04       400001     510000
AE015927_05       500001     610000
AE015927_06       600001     710000
AE015927_07       700001     810000

```

```

AE015927_08       800001     910000
AE015927_09       900001    1010000
AE015927_10      1000001    1110000
AE015927_11      1100001    1210000
AE015927_12      1200001    1310000
AE015927_13      1300001    1410000
AE015927_14      1400001    1510000
AE015927_15      1500001    1610000
AE015927_16      1600001    1710000
AE015927_17      1700001    1810000
AE015927_18      1800001    1910000
AE015927_19      1900001    2010000
AE015927_20      2000001    2110000
AE015927_21      2100001    2210000
AE015927_22      2200001    2310000
AE015927_23      2300001    2410000
AE015927_24      2400001    2510000
AE015927_25      2500001    2610000
AE015927_26      2600001    2710000
AE015927_27      2700001    279251
Continuation (6 of 28) of AE015927 from base 500001 (AE015927 Clostridium tetani E88, cc

Query Match          13.3%; Score 40; DB 15; Length 110000;
Best Local Similarity 51.1%; Pred. No. 0.057;
Matches 94; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
QY      37 AGACTTGAATATTTCTTCATTTGAAATAAAGCAGCTAACTCCCAATTTGCTACTACTAGG 96
DB      21776 AGAATTAATAAGTAGTTATATAAAGGCGATATTTCTTCAATATACCATATGAAA 21835
QY      97 GGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAATAGAAAGTAAACAGAGTG 156
DB      21836 GCAGTATGCTCTATAAATAATATAAATGTTGAAAAATTTATATAAACACTATACCCATA 21895
QY      157 TTGCAACATGCTGTTAATAACCCGACGAACTTCAATCACTATAGCTAGTAGAGTGCA 216
DB      21896 TTATAAATGTTAAATTTAAATTTGCAATCAACCTTTACTATATATAAATAATGGGAT 21955
QY      217 TTCT 220
DB      21956 TTTT 21959

```

```

RESULT 33
AL137018/c
LOCUS              157168 bp    DNA    linear    PRI 18-MAY-2005
DEFINITION        Human DNA sequence from clone RP11-133M9 on chromosome 9q13-21.33
                   Contains the 3' end of the RORB gene for RAR-related orphan
                   receptor B protein (RZRB, ROR-BETA), complete sequence.
ACCESSION         AL137018
VERSION           AL137018.9 GI:14141220
KEYWORDS          HTG; receptor; ROR-BETA; RORB; RZRB.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                  Hominidae; Homo.
REFERENCE         1 (bases 1 to 157168)
AUTHORS          Tracey,A.
TITLE            Direct Submission
JOURNAL          Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
                  Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
                  Clone requests: clonerequest@sanger.ac.uk
                  On May 17, 2001 this sequence version replaced gi:13751249.
                  The following abbreviations are used to associate primary accession
                  numbers given in the feature table with their source databases:
                  Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
                  on the WORMPEP database can be found at
                  http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
                  was generated from part of bacterial clone contigs of human
                  chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
                  Group. Further information can be found at
                  http://www.sanger.ac.uk/HGP/Chr9

```


Nott, A., Nwaokemele, O., Obregon, M., Ochi-Okorie, C., Odeh, E., Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B., Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T., Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quiroz, J., Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S., Rives, C., Rodriguez, F., Rojas, A., Ruiz, S.J., Sana, M., Sanders, W., Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, Y., Sison, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R., Trejos, Z., Umani, K., Vargo, C., Verduzco, D., Villasana, D., Virk, D., Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L., Zhang, Z., Zhou, J., Weinstein, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

124078 124127: gap of 50 bp
 124128 137817: contig of 13690 bp in length
 137818 137867: gap of 50 bp
 137868 150987: contig of 13120 bp in length
 150988 151087: gap of unknown length
 151088 152089: contig of 1002 bp in length
 152090 152189: gap of unknown length
 152190 153599: contig of 1410 bp in length
 153600 153699: gap of unknown length
 153700 154931: contig of 1232 bp in length
 154932 155031: gap of unknown length
 155032 157097: contig of 2066 bp in length
 157098 157197: gap of unknown length
 157198 160933: contig of 3736 bp in length.

FEATURES

Location/Qualifiers

1..160933

/organism="Bos taurus"

/mol_type="genomic DNA"

/db_xref="taxon:9913"

/clone="CH240-265N2"

57472..57521

/estimated_length=50

101522..101727

/estimated_length=206

124078..124127

/estimated_length=50

137818..137867

/estimated_length=50

150988..151087

/estimated_length=unknown

152090..152189

/estimated_length=unknown

153600..153699

/estimated_length=unknown

154932..155031

/estimated_length=unknown

157098..157197

/estimated_length=unknown

ORIGIN

Query Match

Best Local Similarity

Matches

73; Conservative

0; Mismatches

55; Indels

0; Gaps

0;

QY

6 TGACAGGCTAAATCTGCTAGTACCTACGAGGACTTGAATAATTTCTTCAATTCGAAATAA 65

17018 TTACAGGATATAGGATAATCTGCGGAGAGAAACTTGCAGTGCAGCTCCCAATCA 17077

QY

66 ACAGCTAACTCCCAATTTGCTACCTACCTAGGGGAGTTTATCATGAAGAAATGTTTAAATG 125

17078 ATAGCAAGTGCCATACTAAACCTTAACCAAGTGTTACTTTATGTTTAAATCATGATTTCATTT 17137

QY

126 CTTAATTT 133

17138 TGAATAAT 17145

Db

RESULT 35

BX548256

LOCUS

DEFINITION

Zebrafish DNA sequence from clone DKEYP-75B4 in linkage group 1,

complete sequence.

ACCESSION

BX548256

VERSION

BX548256.7

KEYWORDS

HTG.

SOURCE

Danio rerio (zebrafish)

ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

REFERENCE

1 (bases 1 to 181272)

AUTHORS

Barker, D.

TITLE

Direct Submission

NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 NOTE: This is a 'working draft' sequence. It currently
 consists of 10 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information -----
 Center project name: FOKB
 Center clone name: CH240-256G15
 ----- Summary Statistics -----
 Assembly program: Atlas 3.0;
 Consensus quality: 185578 bases at least Q40
 Consensus quality: 186886 bases at least Q30
 Consensus quality: 188093 bases at least Q20
 Estimated insert size: 189246; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 18 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 16328: contig of 16328 bp in length
 * 16329 16378: gap of 50 bp
 * 16379 27546: contig of 11168 bp in length
 * 27547 27596: gap of 50 bp
 * 27597 36718: contig of 9122 bp in length
 * 36719 36768: gap of 50 bp
 * 36769 41738: contig of 4970 bp in length
 * 41739 41798: gap of 60 bp
 * 41799 45763: contig of 3965 bp in length
 * 45764 45813: gap of 50 bp
 * 45814 53873: contig of 8060 bp in length
 * 53874 53973: gap of unknown length
 * 53974 56584: contig of 2611 bp in length
 * 56585 57550: gap of 966 bp
 * 57551 64387: contig of 6837 bp in length
 * 64388 64437: gap of 50 bp
 * 64438 93575: contig of 28138 bp in length
 * 93576 93625: gap of 50 bp
 * 93626 108540: contig of 14915 bp in length
 * 108541 108590: gap of 50 bp
 * 108591 147592: contig of 39002 bp in length
 * 147593 147642: gap of 50 bp
 * 147643 162381: contig of 14739 bp in length
 * 162382 162431: gap of 50 bp
 * 162432 176477: contig of 14046 bp in length
 * 176478 176539: gap of 62 bp
 * 176540 177541: contig of 1002 bp in length
 * 177542 177641: gap of unknown length
 * 177642 185984: contig of 8343 bp in length
 * 185985 187209: gap of 1225 bp
 * 187210 189990: contig of 2781 bp in length
 * 189991 190091: gap of unknown length
 * 190091 191091: contig of 1001 bp in length
 * 191092 191191: gap of unknown length
 * 191192 192411: contig of 1220 bp in length.
 Location/Qualifiers
 1. 192411
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="CH240-256G15"
 16329. 16378
 /estimated_length=50
 27547. 27596
 /estimated_length=50
 36719. 36768
 /estimated_length=50
 41739. 41798

FEATURES
source

gap /estimated_length=60
 45764. 45813
 /estimated_length=50
 53874. 53973
 /estimated_length=unknown
 56585. 57550
 /estimated_length=966
 64388. 64437
 /estimated_length=50
 93576. 93625
 /estimated_length=50
 108541. 108590
 /estimated_length=50
 147593. 147642
 /estimated_length=50
 162382. 162431
 /estimated_length=50
 176478. 176539
 /estimated_length=62
 177542. 177641
 /estimated_length=unknown
 185985. 187209
 /estimated_length=1225
 189991. 190090
 /estimated_length=unknown
 191092. 191191
 /estimated_length=unknown
 ORIGIN
 Query Match 13.3%; Score 40; DB 12; Length 192411;
 Best Local Similarity 57.0%; Pred. No. 0.056;
 Matches 73; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 QY 6 TGACAGGCTAATGCTAAGTCACTACGATGACGACTGCAAAATATCTTCATTGGAATAA 65
 Db 118110 TTACAGGATAATAGGATAATCTGAGGGAAGAAACTTCCCAAGTCACGCTCCCAAAATCA 118051
 QY 66 ACAGCTAACTCCCAAAATGTACTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATG 125
 Db 118050 ATAGCAAGTGCCTAATACCTTAACCTAACAAGTGCTATTATGTTTAATCATGATTCATT 117991
 QY 126 CTTAATTT 133
 Db 117990 TGAATAAT 117983
 RESULT 37
 ABI93599 Plasmodium falciparum 4250 bp mRNA linear INV 28-JUL-2005
 LOCUS Plasmodium falciparum rhoph1/clag8 mRNA for Rhoph1/Clag8, complete
 DEFINITION cds.
 ACCESSION ABI93599 GI:71361251
 VERSION ABI93599.1
 KEYWORDS Plasmodium falciparum (malaria parasite P. falciparum)
 SOURCE Plasmodium falciparum
 ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 REFERENCE 1 Kaneko, O., Yim Lim, B.Y.S., Iriko, H., Ling, I.T., Otsuki, H.,
 Grainger, M., Tsuboi, T., Adams, J.H., Mattei, D., Holder, A.A. and
 Torii, M.
 Apical expression of three Rhoph1/Clag proteins as components of
 the Plasmodium falciparum Rhoph complex
 Mol. Biochem. Parasitol. 143 (1), 20-28 (2005)
 REFERENCE 2 (bases 1 to 4250)
 15953647
 Kaneko, O., Iriko, H., Tsuboi, T. and Torii, M.
 Direct Submission
 Submitted (25-OCT-2004) Osamu Kaneko, Ehime University School of
 Medicine, Department of Molecular Parasitology; Shitsukawa, Toon,
 Ehime, 791-0295, Japan (E-mail: okaneko@ehime-u.ac.jp,
 URL: <http://www.m.ehime-u.ac.jp/school/parasitology/eng/index.html>,
 Tel: 81-89-960-5286, Fax: 81-89-960-5287)

Assembly program: Phrap; version 0.990329
Consensus quality: 140365 bases at least Q40
Consensus quality: 143911 bases at least Q30
Consensus quality: 146396 bases at least Q20
Estimated insert size: 143156; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 33 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1137: contig of 1137 bp in length
1138 1237: gap of unknown length
1238 2283: contig of 1046 bp in length
2284 2383: gap of unknown length
2384 4115: contig of 1732 bp in length
4116 4215: gap of unknown length
4216 5435: contig of 1220 bp in length
5436 5535: gap of unknown length
5536 6746: contig of 1211 bp in length
6747 6847: gap of unknown length
6848 8104: contig of 1258 bp in length
8105 8204: gap of unknown length
8205 9949: contig of 1745 bp in length
9950 10049: gap of unknown length
10050 11148: contig of 1099 bp in length
11149 11248: gap of unknown length
11249 12657: contig of 1409 bp in length
12658 12757: gap of unknown length
12758 15247: contig of 2490 bp in length
15248 15347: gap of unknown length
15348 17176: contig of 1829 bp in length
17177 17277: gap of unknown length
17278 19265: contig of 1989 bp in length
19266 21679: gap of unknown length
21680 21779: gap of unknown length
21780 24238: contig of 2459 bp in length
24239 24338: gap of unknown length
24339 28792: contig of 4454 bp in length
28793 28892: gap of unknown length
28893 31952: contig of 3060 bp in length
31953 32052: gap of unknown length
32053 35524: contig of 3472 bp in length
35525 35624: gap of unknown length
35625 37842: contig of 2218 bp in length
37843 40865: gap of unknown length
40866 40965: gap of unknown length
40966 43954: contig of 2989 bp in length
43955 44054: gap of unknown length
44055 49162: contig of 5108 bp in length
49163 49262: gap of unknown length
49263 53030: contig of 3768 bp in length
53031 53130: gap of unknown length
53131 57558: contig of 4628 bp in length
57559 57858: gap of unknown length
57859 62862: contig of 5004 bp in length
62863 62963: gap of unknown length
62964 69622: contig of 6659 bp in length
69623 77762: gap of unknown length
77763 77861: gap of unknown length
77862 86145: contig of 8284 bp in length
86146 86245: gap of unknown length
86246 96279: contig of 10034 bp in length
96280 96379: gap of unknown length

96380 105100: contig of 8721 bp in length
105101 105200: gap of unknown length
105201 116139: contig of 10939 bp in length
116140 116239: gap of unknown length
116240 129563: contig of 13324 bp in length
129564 129663: gap of unknown length
129664 138778: contig of 9115 bp in length
138779 138878: gap of unknown length
138879 157011: contig of 18133 bp in length.
Location/Qualifiers
1. 157011
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-303C17"
1138. 1237
/estimated_length=unknown
2284. 2383
/estimated_length=unknown
4116. 4215
/estimated_length=unknown
5436. 5535
/estimated_length=unknown
6747. 6846
/estimated_length=unknown
8105. 8204
/estimated_length=unknown
9950. 10049
/estimated_length=unknown
11149. 11248
/estimated_length=unknown
12658. 12757
/estimated_length=unknown
15248. 15347
/estimated_length=unknown
17177. 17276
/estimated_length=unknown
19266. 19365
/estimated_length=unknown
21680. 21779
/estimated_length=unknown
Query Match 13.3%; Score 39.8; DB 12; Length 157011;
Best Local Similarity 46.3%; Pred. No. 0.066;
Matches 131; Conservative 0; Mismatches 152; Indels 0; Gaps 0;
Qy 17 ATGCTAAGTACTAGATGAAGACTTGAATATTTCTTCATTGTAATAAAGCTAATCTC 76
Db 23417 ATGTGAAGGTTATAAGTAAATAATTTAGGGTTTTTTTAAAGATTTATATAATATGATTTT 23476
Qy 77 CCAATTGTACCTAAGTGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCA 136
Db 23477 TGGGAGAAGATTTAAATAAATTTTATTTATATAAATAGAAATGTTAGTTAGTTTAG 23536
Qy 137 AATAAGAAGTAACACAGAGTGTTCACATGCTGTTAAATAACCCGACAACTTCAATCAC 196
Db 23537 AAAAAAGCGTAAATTTGTTATATAAATAAGTTAAGAAATAATTAATAAATAAATAA 23596
Qy 197 TATAGCTGTAGTAGAGTGCATTCGCAAGATCCAGATTAACAGTATTTTGAATGTC 256
Db 23597 AAGAAATAGTTTAAATGTTATTTTGAAGAAGTAAGTAAATTAATAAATTTTGAATA 23656
Qy 257 AATGTTGAACCGACCACTACTAATTTATCTGCTGATTAGGAAGTA 299
Db 23657 AATAATAATTTATGTTTAAAGTAATATGAATATAAATAAATA 23699
RESULT 39
AL929174/c
LOCUS AL929174 161610 bp DNA linear VRT 07-NOV-2003
DEFINITION Zebrafish DNA sequence from clone CH211-117C16 in linkage group 12,
complete sequence.
ACCESSION AL929174


```

TITLE
JOURNAL
REFERENCE
AUTHORS
Worley,K.C.
TITLE
JOURNAL
Submitted (29-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 185657)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (21-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 21, 2002 this sequence version replaced gi:22772818.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
sequence may extend beyond the ends of the clone and there may be
contigs that consist entirely of whole genome shotgun sequence
reads. Both end sequences and whole genome shotgun sequence only
contigs will be indicated in the feature table.
-----
Center: Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: GDPI
Center clone name: CH230-9P18
-----
Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 169494 bases at least Q40
Consensus quality: 172042 bases at least Q30
Consensus quality: 173666 bases at least Q20
Estimated insert size: 184697; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 181626: contig of 181626 bp in length
* 181627 181726: gap of unknown length
* 181727 185657: contig of 3331 bp in length.
-----
Location/Qualifiers
1..185657
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-9P18"
-----
1..1734
/note="wgs_end_extension"
clone_end:T7"
1785..3006
/note="wgs_end_extension"
clone_end:T7"
6683..7563
/note="clone_boundary"
clone_end:T7
site:EcoRI
end_sequence:BH344419"
76575..77257
/note="clone_boundary"
clone_end:Sp6
site:EcoRI
-----
FEATURES
source
1..185657
misc_feature
176639..177753
/note="wgs_end_extension"
clone_end:Sp6"
178823..180024
/note="wgs_end_extension"
clone_end:Sp6"
180462..181626
/note="wgs_end_extension"
clone_end:Sp6"
181627..181726
/estimated_length=unknown
ORIGIN
Query Match 13.3%; Score 39.8; DB 12; Length 185657;
Best Local Similarity 60.7%; Pred. No. 0.066;
Matches 65; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 39 ACTTGAAATATTCATTGCAATTAACAGCTCACTCCCAATTTGTACCTAAGGGG 98
Db 95656 ACTAGAAAGGTAAATTTTAAATAAATAACCTTACATCCACTGTGTTTCAGAAATGAGCAG 95597
QY 99 AGTTATCATCAAGAAATGTTTAAATGCTTAATTTTCAATAAGAG 145
Db 95596 TGTCAATTTCTGACAAATAATACAAATTTGTAATTTTCTCAAAAGAG 95550
-----
RESULT 41
BX005210 Zebrafish DNA sequence from clone DKEY-222F8 in linkage group 4,
complete sequence.
ACCESSION BX005210
VERSION BX005210.10 GI:62719219
KEYWORDS HTG
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
AUTHORS Pelan,S.
TITLE Direct Submission
JOURNAL Submitted (15-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfzfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 17, 2005 this sequence version replaced gi:41016365.
COMMENT
-----
Center: Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfzfsh-help@sanger.ac.uk
-----
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/
Zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.
DKEY-222F8 is from a Zebrafish BAC library
VECTOR: pIndigoBAC-5
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=

```

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC. Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhihong Bao and Sean Eddy, submitted), and those beginning 'drv' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml.

FEATURES

source
1..207949
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clones="DKEY-222P8"
/clone_lib="DanioKey"

ORIGIN

Query Match 13.3%; Score 39.8; DB 11; Length 207949;
Best Local Similarity 50.8%; Pred. No. 0.066;
Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

34 TGAGAGCTGGAATATCTTCATTGTGAATAAACAGCTAACTCCCAATTTGTACCTAACT 93
69897 TAAAACTAGAAATATGCTATCTACTACAAATAGAAATACATCTCACTTTAAGTTGATATC 69956

94 AGGGGAGTTTATCATGATGAAGAAATGTTTAAATGCTTTAAATTTTCAATAAAGAAGTAACAGA 153
69957 CCTGCAGATTAACTGGAATCATGACCTACTATTATTTCCCTTTTGCAGTTCACATTGA 70016

154 GTGTTGCAACATGCTGTTAAATTAACCCGACAACTTCAATCATCATATAGCTGTAGTAGAGT 213
70017 ATAATTTGCGAGGTAACAAATGAATTTTAAACACACAAAACCTCTCAATAGATGGGAAT 70076

214 GCATTCT 220
70077 GCAGACT 70083

RESULT 42

AC126085 Rattus norvegicus clone CH230-14504, WORKING DRAFT SEQUENCE.
AC126085

AC126085.4 GI:25138258

HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 208337)

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, B.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, F., Garza, M.,
Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpthy, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensheewa, L., Loulaseg, H., Lozado, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L. L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajd, D.,
Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 208337)

Worley, K. C.

Direct Submission

Submitted (03-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 208337)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 20, 2002 this sequence version replaced gi:22856016.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GZCO

Center clone name: CH230-14504

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 200812 bases at least Q40

Consensus quality: 202048 bases at least Q30

Estimated insert size: 205307; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submittor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
 source
 1. 208337
 Location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-14504"
 misc_feature
 1. .1006
 /note="wgs_contig"

ORIGIN
 Query Match 13.3%; Score 39.8; DB 12; Length 208337;
 Best Local Similarity 60.7%; Pred. No. 0.066; Mismatches 42; Indels 0; Gaps 0;
 Matches 65; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 39 ACTTGAATATCTTCATTGGAATTAACAGCTAACTCCAAATTTGTACCTAACTAGGGG 98
 Db 184783 ACTGAAAGGTAATTTTAAAAAATAACCTTACATCCACTGTGTTGAGATGAGCAG 184842

Qy 99 AGTTTATCATGAGAAATGTTTAATGCTTAATTTCAATTAAGAG 145
 Db 184843 TGTCAATTTCTGAACAAATTAACAATTGGTAAATTTCTCAAAAGAAAG 184889

RESULT 43
 AC157740 46494 bp DNA linear HTG 25-FEB-2005
 LOCUS
 Xenopus tropicalis clone ISB1-43N23, WORKING DRAFT SEQUENCE, 2
 DEFINITION
 unorderded pieces.
 AC157740
 AC157740.1 GI:60279812
 VERSION
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
 KEYWORDS
 Xenopus tropicalis (Silurana tropicalis)
 SOURCE
 ORGANISM
 Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus; Silurana.
 1 (bases 1 to 46494)
 DOE Joint Genome Institute.
 Unpublished
 JOURNAL
 Unpublished
 2 (bases 1 to 46494)
 DOE Joint Genome Institute.
 Direct Submission
 TITLE
 Submitted (25-FEB-2005) Production Genomics Facility, DOE Joint
 JOURNAL
 Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
 94598-1698, USA
 -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: <http://www.jgi.doe.gov>

 Project Information
 Center Project Name: 2789798
 Center clone name: ISB-43N23

 Summary Statistics
 Consensus quality: 45777 bases at least Q40
 Consensus quality: 46121 bases at least Q30
 Consensus quality: 46222 bases at least Q20
 Estimated insert size: 75000; agarose-fp estimation
 Quality coverage: 46394; sum-of-contigs estimation
 Quality coverage: 8.24 in Q20 bases; agarose-fp estimation

Quality coverage: 13.32 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

FEATURES
 source
 1. 17494; contig of 17494 bp in length
 17495 17594; gap of unknown length
 17595 46494; contig of 28900 bp in length.
 Location/Qualifiers
 1. .46494
 /organism="Xenopus tropicalis"
 /mol_type="genomic DNA"
 /db_xref="taxon:8364"
 /clone="ISB1-43N23"
 /clone_lib="ISB Xenopus tropicalis BAC library"
 17495. .17594
 /estimated_length=unknown
 gap
 ORIGIN
 Query Match 13.2%; Score 39.6; DB 12; Length 46494;
 Best Local Similarity 49.5%; Pred. No. 0.078; Mismatches 102; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy 79 AAATTGTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTCAA 138
 Db 13333 AAATTGTATCATTCATGAATAATAATTAATGATAGATATTTCTAAGGATTTAATGAA 13392

Qy 139 TAAGAAGTAACAGAGTGTTCACATGCTGTTAAATAACCCGACAACTTCAATCACTA 198
 Db 13393 GAACCTTTTGCCAAAGTGTGCTCAATTTTAAAGAAAGAGAGAGACTGTAACCTGTT 13452

Qy 199 TAGCTGTAGTAGTGCAATTCGCAAGATCCCGAGATACCAAGTATTTTGGAAATGCAA 258
 Db 13453 GCACCCCTTTATTGTGGAATAAAGATGCACAGAAAGTAAACATTAAACATGCAATCTGC 13512

Qy 259 TGTGTAACCGACCATCACTAATATCT 284
 Db 13513 CATCCACAGACCTAAATAAATGTAT 13538

RESULT 44
 AC133155.2
 WPCOMMENT
 Sequence split into 4 fragments LOCUS AC133155 Accession AC133155
 Fragment Name Begin End
 AC133155.0 1 110000
 AC133155.1 100001 210000
 AC133155.2 200001 310000
 AC133155.3 300001 398299
 Continuation (3 of 4) of AC133155 from base 200001 (AC133155 Mus musculus clone RP24-418)

Query Match 13.2%; Score 39.6; DB 12; Length 110000;
 Best Local Similarity 52.4%; Pred. No. 0.077; Mismatches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 45 AATATTCTTCATTGGAATAAAGAGTAACTCCCAATTTGACCTAATAGGGAGTTTA 104
 Db 37295 AACTTTCTTTTATTGAAATAAAGAGATGGGTGGCTAGGAAAGGGGTGGA 37354

Qy 105 TCATGGAAGAATGTTTAAATGCTTAATTTCAATATAGAGTACCAAGAGTGTTCACACA 164
 Db 37355 TCTAGGAGAGTCTGGGAAGGGGTGAATGTGATAAAAAATGTTATGAGAGGTTTCAAAA 37414

Qy 165 TGCTGTTAAATAACCCGACAACTTCAATCACTATAGCTGTAGTAG 210
 Db 37415 AACTGATAATATGAGGAGGGAACAAATAAATAATGCAAGTTTATG 37460

RESULT 45

Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 154673)

REFERENCE AUTHORS

Anderson, M., Musbaum, C., Lander, E., Abouelleil, A., Allen, N., Bloom, T., Boguslavsky, L., Bouckghalter, B., Camarata, J., Chang, J., Chepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearrellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Katat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE JOURNAL

Submitted (02-NOV-2005) Broad Institute of MIT and Harvard, 320

Charles Street, Cambridge, MA 02141, USA

On Nov 2, 2005 this sequence version replaced gi:77798120.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Broad Institute of MIT and Harvard

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L33519

Center clone name: 380_A_15

FEATURES

source

1. 154673

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="15"

/map="15"

/clone="RP24-380A15"

/clone_lib="RPC1-24 Male Mouse BAC"

1381..1419

/rpt_family="AT-rich"

complement(3059..3399)

/rpt_family="ORR1D"

3644..3664

/rpt_family="AT-rich"

3676..3807

/rpt_family="B1_MM"

4024..4046

/rpt_family="B4"

4047..4175

/rpt_family="B1_MM"

4176..4199

/rpt_family="B4"

4507..4529

/rpt_family="(T)n"

complement(4991..6108)

/rpt_family="L1_MM"

complement(6102..7005)

/rpt_family="L1"

complement(8198..8545)

/rpt_family="Lx2"

9547..9740

/rpt_family="B3"

9743..9796

repeat_region
9827..9847
/rpt_family="AT-rich"
complement(10690..11695)
repeat_region
11693..12379
/rpt_family="Lx4"
repeat_region
14806..14859
/rpt_family="Lx4"
repeat_region
14913..15020
/rpt_family="B3A"
complement(15036..15159)
repeat_region
15222..15337
/rpt_family="Lx2"
repeat_region
15338..15554
/rpt_family="GA-rich"
repeat_region
15338..15554
/rpt_family="L1"
complement(15595..15717)
repeat_region
15721..15812
/rpt_family="RMER17A2"
repeat_region
/rpt_family="L1"
complement(16321..16436)
repeat_region
/rpt_family="MIR"
17416..17428
/rpt_family="B4A"
repeat_region
17429..17509
/rpt_family="B1_MM"
repeat_region
17525..17562
/rpt_family="CA)n"
repeat_region
17636..17712
/rpt_family="(TA)n"
repeat_region
18802..18964
/rpt_family="(TATAA)n"
repeat_region
19344..19586
/rpt_family="B3"
complement(19793..19865)
repeat_region
19936..19957
/rpt_family="MER99"
repeat_region
/rpt_family="AT-rich"
20138..20231
/rpt_family="CT-rich"
complement(20258..20551)
repeat_region
/rpt_family="Lx8"
complement(21400..21514)
repeat_region
/rpt_family="PB1"
21666..21701
/rpt_family="(TG)n"
repeat_region
22616..22781
/rpt_family="RSINE1"
complement(22851..22961)
repeat_region
/rpt_family="Lx9"
23041..23107
/rpt_family="L1MC3"
repeat_region
23123..23234
/rpt_family="PB1D10"
repeat_region
23243..23346
/rpt_family="L1MC3"
repeat_region
23363..23614
/rpt_family="L1MC3"
repeat_region
23615..23652
/rpt_family="(ATGTG)n"
repeat_region
23653..23769
/rpt_family="L1MC3"
repeat_region
25065..25125

Query Match 13.2%; Score 39.6; DB 6; Length 154673;

Best Local Similarity 52.4%; Pred.No.0.077;

Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 45 AATATTTCTTCAATTTGAATTAACAGTAACTCCCAATTTGTACCTACTAGGGGAGTTTA 104

Db 87286 AACTTTCTTTTATTGTAATAAAAAAGGACACAAAGATGGGTGGGTAGGAAAGGGGGTGA 87345


```

Query Match      13.2%; Score 39.6; DB 6; Length 193188;
Best Local Similarity 50.5%; Pred. No. 0.077;
Matches 96; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 40 CTTGAATATTTCTTCATTGAAATTAACAGCTAACTCCCAATTTGTACCTAACTAGGGGA 99
DB 75475 CTAGATAGCTATGAATTTGTAATAGAGATTCACACAGACACAGCAATTAAGCAA 75416
QY 100 GTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAATTAAGAAAGTAAACAGAGTGTG 159
DB 75415 TTTAAAAAGAGAAAAAATATGAGTTCTAGTTTTCAGAAAGAAAGAGATTGTAATTTAT 75356
QY 160 CAACATGCTGTTAAATAACCGGCAAACTCACTACTATAGCTGTAGTAGTGCAATTC 219
DB 75355 TAACTATTTTATATTAAAGAGCAAAATTTCAATTTATTGTAACAGCAGTAGCCTGTTTGA 75296
QY 220 TGCAGGATC 229
DB 75295 CAGAAATATC 75286

RESULT 49
AC084288
LOCUS AC084288 200393 bp DNA linear ROD 31-DEC-2002
DEFINITION Mus musculus Strain C57BL6/J Chromosome 2 BAC, RP23-108H5, complete
sequence.
ACCESSION AC084288
VERSION AC084288.9 GI:27436766
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 200393)
MONTGOMERY, K.T., GRILLS, G., HAN, J., CHIU, D., DECKER, J., FUSINA, M.,
GOLTZ, J., HAIDER, A., HALL, L., IOSHIKHES, I.P., LEE, E., LONG, J.,
PERERA, A., SHIM, C., THOMAS, E. and KUCHERLAPATI, R.
High Throughput Mouse Sequencing
Unpublished
2 (bases 1 to 200393)
MONTGOMERY, K.T., GRILLS, G., HAN, J., CHIU, D., DECKER, J., FUSINA, M.,
GOLTZ, J., HAIDER, A., HALL, L., IOSHIKHES, I.P., LEE, E., LONG, J.,
PERERA, A., SHIM, C., THOMAS, E. and KUCHERLAPATI, R.
Direct Submission
Submitted (21-OCT-2000) Department of Molecular Genetics, Albert
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
Bronx, NY 10461, USA
3 (bases 1 to 200393)
MONTGOMERY, K.T., GRILLS, G., HAN, J., CHIU, D., DECKER, J., FUSINA, M.,
GOLTZ, J., HAIDER, A., HALL, L., IOSHIKHES, I.P., LEE, E., LONG, J.,
PERERA, A., SHIM, C., THOMAS, E. and KUCHERLAPATI, R.
Direct Submission
Submitted (29-JUL-2002) Harvard Partners Center for Genetics and
Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA
02139, USA
4 (bases 1 to 200393)
MONTGOMERY, K.T., GRILLS, G., HAN, J., CHIU, D., DECKER, J., FUSINA, M.,
GOLTZ, J., HAIDER, A., HALL, L., IOSHIKHES, I.P., LEE, E., LONG, J.,
PERERA, A., SHIM, C., THOMAS, E., BROWN, W.A. and KUCHERLAPATI, R.
Direct Submission
Submitted (31-DEC-2002) Harvard Partners Center for Genetics and
Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA
02139, USA
COMMENT On Dec 31, 2002 this sequence version replaced gi:22002114.
-----Genome Center:
Center: Harvard Partners Genome Center
Center Code: HPGC
Web site: http://www.hpcgg.org/Sequence/mouse.html
Contact: hpgc@mendel.mgh.harvard.edu

CLONE LENGTH: This sequence represents the entire insert of this
clone unless otherwise noted. If there are overlapping clones, the
overlaps are noted in the beginning and end of the Features
listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550).
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Regions of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to EST
and cDNA sequences in Unigene. Genes demonstrate at least two exons
flanked by consensus splice sites that maintain sequence continuity
across the splice junctions. Sequences that are not identical
matches are annotated as similar.

SEQUENCING READ COVERAGE: Attempts are made to complete double
stranded sequence for all regions. All sequence is completed to a
standard of coverage with a minimum of 3 reads with no ambiguities.
If the sequence coverage for a region does not meet this standard,
it is indicated in the annotation as Low Coverage. Low coverage
linkages are verified by PCR product size verification or
verification of forward and reverse reads from clones which span
the low coverage area.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated average error rate is less than 1 per 10,000
bases using the Consed quality parameters. Regions that do not
meet this requirement are annotated as Low Quality.

-----Summary Statistics
Center project name: ADJ
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 100%
Assembly program: Phrap version 0.990329
Contig length: 200393
Fraction of Phrap value < 40: 0.001248
Error Rate in Consed: 0.01 per 10,000 bases
Number of N's in consensus: 0

----- Distribution of Quality < 40 Bases:
# bases
1000
900
800
700
600
500
400
300
200
100
0
5 10 15 20 25 30 35 40
Phrap Value Range

FEATURES
source Location/Qualifiers
1. 200393
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-108H5"
81.159
repeat_region /rpt_family="(TG)n"
160.215
misc_feature /standard_name="Low coverage"
/notes="Low Coverage"
1156.1228
repeat_region /rpt_family="MIR3"
8019.8356
repeat_region /rpt_family="(TCTA)n"
complement(10415..10518)

```

| | |
|---------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| repeat_region | /rpt_family="PBID10" |
| misc_feature | 10870..10898 /rpt_family="AT_rich" 12332..12455 /standard_name="low coverage" /note="Low coverage, Unsure Length, CT Repeat, size confirmed by PCR" |
| repeat_region | 12344..12422 /rpt_family="(TC)n" |
| repeat_region | 15917..15956 /rpt_family="(TTTTG)n" |
| repeat_region | complement(16855..16918) /rpt_family="MER116" |
| repeat_region | 18562..18696 /rpt_family="MERVL" |
| repeat_region | complement(19493..19642) /rpt_family="MER5A" |
| repeat_region | 19697..20093 /rpt_family="MTC" |
| repeat_region | complement(20426..20621) /rpt_family="MIR" |
| repeat_region | 22447..22505 /rpt_family="(CA)n" |
| repeat_region | complement(22557..22943) /rpt_family="MTD" |
| repeat_region | complement(23181..23330) /rpt_family="MIR" |
| repeat_region | 25775..25911 /rpt_family="MER5A" |
| repeat_region | 27903..28185 /rpt_family="B4A" |
| repeat_region | 28219..28299 /rpt_family="(CA)n" |
| repeat_region | 28418..28612 /rpt_family="B3" |
| repeat_region | 29435..29535 /rpt_family="PBID10" |
| repeat_region | 29563..29605 /rpt_family="Charlie5" |
| repeat_region | complement(30231..30457) /rpt_family="URR1A" |
| repeat_region | complement(31394..31501) /rpt_family="MIR" |
| repeat_region | 31665..31689 /rpt_family="(TA)n" |
| repeat_region | complement(32793..32844) /rpt_family="ID5" |
| repeat_region | complement(32845..33071) /rpt_family="URR1A" |
| repeat_region | complement(33072..33128) /rpt_family="ID5" |
| repeat_region | complement(33129..33259) /rpt_family="B1F" |
| repeat_region | complement(33260..33387) /rpt_family="MER5A" |
| STS | 33778..33861 /standard_name="J05163" |
| repeat_region | 34411..34454 /rpt_family="(TTTTG)n" |
| repeat_region | 35355..35390 /rpt_family="A-rich" |
| repeat_region | 35447..35573 /rpt_family="RMER17B" |
| STS | 36263..36403 /standard_name="AW146109" |
| repeat_region | 36629..36654 /rpt_family="(TTTTG)n" |
| repeat_region | complement(36710..36777) /rpt_family="ID_B1" |
| repeat_region | 36783..36834 /rpt_family="(TG)n" |
| repeat_region | complement(36886..37134) /rpt_family="B4A" |

| repeat_region | complement(37135..37195) |
|-----------------------|--------------------------------------------------------------------------|
| repeat_region | /rpt_family="ID_B1" |
| repeat_region | 37440..37643 |
| repeat_region | /rpt_family="ID_B1" |
| Query Match | 13.2%; Score 39.6; DB 6; Length 200393; |
| Best Local Similarity | 50.5%; Pred. No. 0.077; |
| Matches | 96; Conservative 0; Mismatches 94; Indels 0; Gaps 0; |
| Qy | 40 CTTGAAATATCTTCATTGTAATAAAGCAGTAACCTCCCAATTTGCTACTAAGGGGA 99 |
| Db | 10759 CTAGAATAGCTATGAATTTGTGAATAGAAATTCACACACAGACACAAATTAAGCAA 10818 |
| Qy | 100 GTTTATCATGAAGAAATGTTAAATGCTTAAATTTTCAAATAGAAGTAACACAGAGTGTG 159 |
| Db | 10819 TTTAAAGAGAAAATAATGAGTTCGTAGTTTCGAAGAAGAAGATTGTAAATTAAT 10878 |
| Qy | 160 CAACATGCTGTTAAATAAACCACAAAACCTTCAATCACTATAGCTGTAGTAGAGTGCATTC 219 |
| Db | 10879 TAACATATTTTATATTTAAAGAGCAAATTTTCATTTATTGTAACAGCAGTAGCGCTTTGA 10938 |
| Qy | 220 TGCAAGGATC 229 |
| Db | 10939 CAGANAATATC 10948 |
| RESULT 50 | |
| EX908768/c | |
| LOCUS | EX908768 229628 bp DNA linear HTG 27-MAR-2004 |
| DEFINITION | Mus musculus chromosome 2 clone RP24-139J5, 5 unordered pieces. |
| ACCESSION | Mus908768 |
| VERSION | EX908768.3 GI:45772194 |
| KEYWORDS | HTG; HTGS; PHASE1; HTGS CANCELLED. |
| SOURCE | Mus musculus (house mouse) |
| ORGANISM | Mus musculus |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| | Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; |
| | Sciurognathi; Muridae; Murinae; Mus. |

| TITLE | JOURNAL | COMMENT |
|--------------------------------------------------------------|---------------------------------|-----------------------------------|
| 1. The Role of the Teacher in the Classroom | Journal of Educational Research | 1980, Vol. 83, No. 1, pp. 1-10 |
| 2. The Impact of Technology on Education | Journal of Educational Research | 1980, Vol. 83, No. 2, pp. 11-20 |
| 3. The Importance of Parental Involvement | Journal of Educational Research | 1980, Vol. 83, No. 3, pp. 21-30 |
| 4. The Effect of Teacher Expectations on Student Achievement | Journal of Educational Research | 1980, Vol. 83, No. 4, pp. 31-40 |
| 5. The Role of the School in the Community | Journal of Educational Research | 1980, Vol. 83, No. 5, pp. 41-50 |
| 6. The Impact of Teacher Education on Student Achievement | Journal of Educational Research | 1980, Vol. 83, No. 6, pp. 51-60 |
| 7. The Importance of Teacher Evaluation | Journal of Educational Research | 1980, Vol. 83, No. 7, pp. 61-70 |
| 8. The Effect of Teacher Salary on Student Achievement | Journal of Educational Research | 1980, Vol. 83, No. 8, pp. 71-80 |
| 9. The Role of the School in the Community | Journal of Educational Research | 1980, Vol. 83, No. 9, pp. 81-90 |
| 10. The Impact of Teacher Education on Student Achievement | Journal of Educational Research | 1980, Vol. 83, No. 10, pp. 91-100 |

* as soon as it is available and the accession number will
* be preserved.

1 154697: contig of 154697 bp in length
* 154698 154797: gap of 100 bp
* 154798 157100: contig of 2303 bp in length
* 157101 157200: gap of 100 bp
* 157201 211216: contig of 54016 bp in length
* 211217 211316: gap of 100 bp
* 211317 220309: contig of 8993 bp in length
* 220310 220409: gap of 100 bp
* 220410 229628: contig of 9219 bp in length.

FEATURES

source

1. .229628
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP24-139J5"
/clone_lib="RPCI-24"

misc_feature

1. .154697
/note="assembly_fragment:03328
clone_end:SP6
vector_side:left"

misc_feature

154798..157100
/note="assembly_fragment:00059"

misc_feature

157201..211216
/note="assembly_fragment:03367
fragment_chain:1"

misc_feature

211317..220309
/note="assembly_fragment:00391
fragment_chain:1"

misc_feature

220410..229628
/note="assembly_fragment:00388
fragment_chain:1
clone_end:T7
vector_side:right"

ORIGIN

Query Match 13.2%; Score 39.6; DB 12; Length 229628;
Best Local Similarity 50.5%; Pred. No. 0.077;
Matches 96, Conservative 0; Mismatches 94; Indels 0; Gaps 0;
Qy 40 CTTGAAATATCTTCATTGAAATPAAACAGCTAACTCCCAAAATGTACCTAACTAGGGGA 99
Db 80862 CTAGATAGCTATGCAATTTGTAATAGAGAATTCACACACAGACACAAATTAAGCAA 80803
Qy 100 GTTATCATGAGAAATGTTAAATGCTTAATTTTCAATTAAGAAAGTAACAGAGTGTG 159
Db 80802 TTTAAAGAGAAAAAATATGAGTCTAGTTTGCAGAAAGAAAGATTGTAATTAAT 80743
Qy 160 CAACATGCTGTTAAATAACCCGACAACTTCAATCACTATAGCTGTAGTAGAGTGCATTC 219
Db 80742 TAACATTTTATATTTAAGAGCAATTTTCATTTATGTAACAGCAGTAGCCCTGTTGA 80683
Qy 220 TGCAAGGATC 229
Db 80682 CAGAAATATC 80673

Search completed: July 17, 2006, 22:34:45
Job time : 2391 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 21:41:41 ; Search time 386 Seconds
(without alignments)
5418.846 Million cell updates/sec

Title: SEQ1-96535C

Perfect score: 300

Sequence: 1.9gcattgacaggctaaatgc.....atctgctgattaggaagtat 300

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : N_Geneseq_8.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005as.*
- 15: Geneseqn2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 299.6 | 100.0 | 110000 | 12 | ADO79173 KLF12 gen |
| 2 | 41.8 | 13.9 | 2000 | 11 | ACL35363 Rice stre |
| 3 | 39.4 | 13.1 | 60461 | 14 | AED18414 Fibrotic |
| 4 | 38.8 | 12.9 | 295096 | 11 | ACN44068 Mouse gen |
| 5 | 38.6 | 12.9 | 2224 | 5 | ABV27665 Human pro |
| 6 | 38.6 | 12.9 | 2224 | 5 | ABV21841 Human pro |
| 7 | 38.6 | 12.9 | 2224 | 5 | ABV24590 Human pro |
| 8 | 38.6 | 12.9 | 2224 | 5 | ABV28008 Human pro |
| 9 | 38.6 | 12.9 | 2224 | 5 | ABV23184 Human pro |
| 10 | 38.6 | 12.9 | 2224 | 5 | ABV24513 Human pro |
| 11 | 38.6 | 12.9 | 2224 | 5 | ABV24601 Human pro |
| 12 | 38.6 | 12.9 | 2224 | 5 | ABV29023 Human pro |
| 13 | 38.6 | 12.9 | 2224 | 5 | ABV28014 Human pro |
| 14 | 38.6 | 12.9 | 2224 | 5 | ABV22169 Human pro |
| 15 | 38.6 | 12.9 | 2224 | 5 | ABV22175 Human pro |
| 16 | 38.6 | 12.9 | 2224 | 5 | ABV25057 Human pro |
| 17 | 37.2 | 12.4 | 14920 | 6 | ABN80147 Human che |
| 18 | 37 | 12.3 | 6857 | 11 | ADW53569 Human 98P |

| | | | | | | |
|------|------|------|--------|----|-------------|--------------------|
| c 19 | 37 | 12.3 | 6857 | 11 | ADW53577 | Adw53577 Human 98P |
| c 20 | 37 | 12.3 | 6857 | 11 | ADW53581 | Adw53581 Human 98P |
| c 21 | 37 | 12.3 | 6857 | 11 | ADW53696 | Adw53696 Partial h |
| c 22 | 37 | 12.3 | 6857 | 11 | ADW53561 | Adw53561 Human 98P |
| c 23 | 37 | 12.3 | 6857 | 11 | ADW53583 | Adw53583 Human 98P |
| c 24 | 37 | 12.3 | 6857 | 11 | ADW53579 | Adw53579 Human 98P |
| c 25 | 37 | 12.3 | 6857 | 11 | ADW53573 | Adw53573 Human 98P |
| c 26 | 37 | 12.3 | 6857 | 11 | ADW53575 | Adw53575 Human 98P |
| c 27 | 37 | 12.3 | 6857 | 11 | ADW53571 | Adw53571 Human 98P |
| c 28 | 37 | 12.3 | 6857 | 11 | ADW53698 | Adw53698 Partial h |
| c 29 | 37 | 12.3 | 6857 | 11 | ADW53585 | Adw53585 Human 98P |
| c 30 | 37 | 12.3 | 6857 | 11 | ADW53559 | Adw53559 Human 98P |
| c 31 | 37 | 12.3 | 6857 | 11 | ADW53563 | Adw53563 Human 98P |
| c 32 | 37 | 12.3 | 6857 | 11 | ADW53525 | Adw53525 Human 98P |
| c 33 | 37 | 12.3 | 6857 | 11 | ADW53565 | Adw53565 Human 98P |
| c 34 | 37 | 12.3 | 6857 | 11 | ADW53567 | Adw53567 Human 98P |
| c 35 | 37 | 12.3 | 6857 | 13 | ADR66608 | Adr66608 Human pro |
| c 36 | 37 | 12.3 | 6857 | 13 | ADR66266 | Adr66266 Human pro |
| c 37 | 36.8 | 12.3 | 6857 | 14 | AEC20351 | Aec20351 cDNA enco |
| c 38 | 36.8 | 12.3 | 2272 | 4 | AAH18058 | Aah18058 Human cdn |
| c 39 | 36.8 | 12.3 | 2533 | 4 | AAK84697 | Aak84697 Human imm |
| c 40 | 36.8 | 12.3 | 2533 | 4 | AAK84698 | Aak84698 Human imm |
| c 41 | 36.6 | 12.2 | 82761 | 14 | AED18524 | Aed18524 Fibrotic |
| c 42 | 36.4 | 12.1 | 320 | 13 | ACF86506 | Acf86506 Human SIR |
| c 43 | 36.4 | 12.1 | 392 | 13 | ACF88431 | Acf88431 Human SIR |
| c 44 | 36.4 | 12.1 | 2374 | 15 | AEF55932 | Aef55932 Human gen |
| c 45 | 36.4 | 12.1 | 2374 | 15 | AEF55696 | Aef55696 Human gen |
| c 46 | 36.2 | 12.1 | 2000 | 11 | ACL35363 | ACL35363 Rice stre |
| c 47 | 36.2 | 12.1 | 9087 | 6 | ABK31277 | Abk31277 Signal tr |
| c 48 | 36.2 | 12.1 | 9087 | 6 | ABL70238 | AbL70238 Chemical |
| c 49 | 36.2 | 12.1 | 9087 | 6 | AAS61181 | Aas61181 Human gen |
| c 50 | 36.2 | 12.1 | 17738 | 6 | ABL33538 | AbL33538 Human imm |
| c 51 | 36.2 | 12.1 | 37992 | 13 | ABD33572 | Abd33572 Murine ca |
| c 52 | 36 | 12.0 | 6020 | 6 | ABL32199 | AbL32199 Human imm |
| c 53 | 36 | 12.0 | 6020 | 6 | ABL49308 | AbL49308 Human pol |
| c 54 | 35.8 | 11.9 | 2310 | 8 | ABZ10134 | AbZ10134 Haematopo |
| c 55 | 35.8 | 11.9 | 5310 | 10 | ADE84160 | Ade84160 Human lym |
| c 56 | 35.8 | 11.9 | 7588 | 4 | AAS45483 | Aas45483 Chemical |
| c 57 | 35.8 | 11.9 | 7588 | 6 | ABL34012 | AbL34012 Human imm |
| c 58 | 35.8 | 11.9 | 7588 | 6 | ABK28409 | Abk28409 DNA trans |
| c 59 | 35.6 | 11.9 | 2000 | 11 | ACL35887 | ACL35887 Rice stre |
| c 60 | 35.6 | 11.9 | 130263 | 6 | ABK83573 | AbK83573 Human cdn |
| c 61 | 35.6 | 11.9 | 277616 | 13 | ABD32602 | Abd32602 Human can |
| c 62 | 35.4 | 11.8 | 2381 | 5 | AAS15800 | Aas15800 Human Six |
| c 63 | 35.4 | 11.8 | 3754 | 11 | ADW53521 | Adw53521 Human 98P |
| c 64 | 35.4 | 11.8 | 3754 | 11 | ADW53684 | Adw53684 Partial h |
| c 65 | 35.4 | 11.8 | 3754 | 11 | ADW53686 | Adw53686 Partial h |
| c 66 | 35.4 | 11.8 | 4329 | 5 | AAS15801 | Aas15801 Human ORF |
| c 67 | 35.4 | 11.8 | 23054 | 4 | AAK84676 | Aak84676 Human imm |
| c 68 | 35.4 | 11.8 | 37973 | 6 | ABL34196 | AbL34196 Human imm |
| c 69 | 35.4 | 11.8 | 99918 | 13 | ABD33009 | Abd33009 Human can |
| c 70 | 35.2 | 11.7 | 4220 | 6 | ABN81677 | Abn81677 Human CDK |
| c 71 | 35.2 | 11.7 | 4220 | 6 | ABN81676 | Abn81676 Human CDK |
| c 72 | 35.2 | 11.7 | 4233 | 6 | AAS61454 | Aas61454 Human gen |
| c 73 | 35.2 | 11.7 | 4233 | 8 | ABZ10093 | AbZ10093 Haematopo |
| c 74 | 35.2 | 11.7 | 4233 | 8 | ABZ09989 | Abz09989 Haematopo |
| c 75 | 35.2 | 11.7 | 4233 | 10 | ADBE84073 | Ade84073 Human lym |
| c 76 | 35.2 | 11.7 | 110000 | 4 | AAK95240_02 | Continuation (3 of |
| c 77 | 35.2 | 11.7 | 110000 | 4 | AAK96733_02 | Continuation (3 of |
| c 78 | 35.2 | 11.7 | 110000 | 6 | ABT00010_02 | Continuation (3 of |
| c 79 | 35.2 | 11.7 | 110000 | 6 | ABT01503_02 | Continuation (3 of |
| c 80 | 35.2 | 11.7 | 110000 | 12 | ADW70291_02 | Continuation (3 of |
| c 81 | 35.2 | 11.7 | 110000 | 12 | ADH77486_02 | Continuation (3 of |
| c 82 | 35 | 11.7 | 476 | 5 | ABV06229 | Abv06229 Human pro |
| c 83 | 35 | 11.7 | 2000 | 11 | ACL37108 | ACL37108 Rice stre |
| c 84 | 35 | 11.7 | 6912 | 6 | ABK28372 | AbK28372 DNA trans |
| c 85 | 35 | 11.7 | 61020 | 4 | AAS46788 | Aas46788 Tumour su |
| c 86 | 35 | 11.7 | 110000 | 14 | AEE04879_3 | Continuation (4 of |
| c 87 | 34.8 | 11.6 | 1056 | 3 | AAC59528 | Aac59528 Human sec |
| c 88 | 34.8 | 11.6 | 2646 | 15 | AEF55793 | Aef55793 Human gen |
| c 89 | 34.8 | 11.6 | 18133 | 6 | ABK40017 | Abk40017 Human che |
| c 90 | 34.8 | 11.6 | 18133 | 6 | ABL32940 | AbL32940 Human imm |
| c 91 | 34.8 | 11.6 | 110000 | 5 | AAI61373_4 | Continuation (5 of |

QY 78 CAAATTGTACCTAAGGAGGTTTATCATGAGAAATGTTTAAATGCTTTAAATTTTCAA 137
 DB 27784 CAGAAGTTATTAAAGTAAAGAGGCAACAAAGAGAGAAATGATTAACAAATATGTAA 27725
 QY 138 ATAGAGTACACAGAG 154
 DB 27724 CTAATAAATAAGAGAG 27708
 RESULT 4
 ACN44068
 ID ACN44068 standard; DNA; 295096 BP.
 XX ACN44068;
 AC
 DT 18-NOV-2004 (first entry)
 XX Mouse genomic sequence MCG13636.
 DB
 DE Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
 XX
 KW Mus musculus.
 XX
 OS WO2003073826-A2.
 XX
 PN 12-SEP-2003.
 XX
 PD 28-FEB-2003; 2003WO-US006235.
 XX
 PE 01-MAR-2002; 2002US-00087192.
 XX
 PF (SAGR-) SAGRES DISCOVERY.
 XX
 PG Morris DW;
 XX
 PH WPI; 2003-328604/31.
 XX
 PI Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
 PT comprises a nucleotide sequence.
 PT
 PS Claim 1; SEQ ID NO 331; Opp; English.
 XX
 CC The present invention relates to novel DNA and protein sequences which
 CC are associated with carcinomas. The sequences are useful for: (i) for
 CC screening drug candidates; (ii) for screening of bioactive agent capable
 CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
 CC determining Carcinoma Associated (CA) gene copy number. In addition, the
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
 CC carcinoma including lymphoma. The present sequence is one such CA coding
 CC sequence. Note: This patent is an equivalent to basic patent
 CC US2002182586A1, for which no sequence data was published
 CC
 SQ Sequence 295096 BP; 75726 A; 62659 C; 64620 G; 85924 T; 0 U; 6167 Other;
 Query Match 12.9%; Score 38.8; DB 11; Length 295096;
 Best Local Similarity 48.9%; Pred. No. 4.3;
 Matches 132; Conservative 0; Mismatches 137; Indels 1; Gaps 1;
 QY 1 GGCATTGACAGGCTAAATGCTAGTACTAGTACGAGAGCTTGAATATTTCTTCTTGA 60
 DB 175844 GCCATTGCTGAAGACTCTGCTTGGAGGCTAGATGAAGAGATTCATATTATTCCTTGA 175903
 QY 61 AATAAAGAGCTAACTCCCAATTTCTACTAGGAGGAGTTTATCATGAGAAATGTTT 120
 DB 175904 AAGAAGAGCTCAAAACAGCTCTGCTATTAATTTTGTGTTTATTCAGTTCTACTCTTA 175963
 QY 121 AAATGCTTAATTTTCAAAATAAGAGTAACCAAGAGTGTTCACATGCTGT-TAAATAACC 179

DB 175964 TGAGGAATGTTTAAATAAAGAGGAGTAAGCTGAGAAAGGAAATAACAAAATATATGCT 176023
 QY 180 CGACAAACTTCAATCACTATATAGCTGTAGTAGAGTGCAATTCGCAAGGATCCCAAGAGTAAC 239
 DB 176024 CAAGTATTAAAGGGGCATCAAGAAAGTAGAATGGAGCTGAATCCTATGTTCAGGAGTTAA 176083
 QY 240 CAGTATTGGAATGCAATGTTGAACCGA 269
 DB 176084 CAAATAAGGAGGTGGTAATCTTGAGGCAA 176113
 RESULT 5
 ABV27665
 ID ABV27665 standard; cDNA; 2224 BP.
 XX AC
 AC ABV27665;
 XX
 DT 16-SEP-2002 (first entry)
 XX Human prostate expression marker cDNA 27656.
 DE
 DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PE 20-FEB-2001; 2001WO-US005171.
 XX
 PF 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX
 (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 WPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 5665; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;
 Query Match 12.9%; Score 38.6; DB 5; Length 2224;
 Best Local Similarity 56.5%; Pred. No. 1.3;
 Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
 QY 41 TTGAAATATTCTTCATTGTAATAAAGAGTAACTCCCAATTTGTACCTAATAGGGAG 100

| | | | |
|----|------|--------------------------------------------------------------------|------|
| Db | 2035 | TTGAAATTATTTTAAAGCCTTAAAAATTCCTGCTGGTTTGCAGAAATTCGTATATAACTATATATC | 2094 |
| Qy | 101 | TTTATCATGAA -GAAATGTTTAAATGCTTAAATTTTCAAATAAGAAAGTAACACAGAGTGTG | 159 |
| Db | 2095 | TATATTTAACTCTGTAAGTATTTTGGCTGATTAAGAAATATAAAGTAAACAACAGGGTTA | 2154 |
| Qy | 160 | CAACATCCTCTGTTAAATAACCCGACAAACTTCCAATCACTATA | 200 |
| Db | 2155 | CAGTTGTGCAAAAACAAAACAAAACATACCTAAATTTACAAAT | 2195 |

| | |
|----------|-------------------------------------------------------------------------|
| RESULT 6 | |
| ABV21841 | ABV21841 standard; cDNA; 2224 BP. |
| XX | |
| XX | ABV21841; |
| XX | |
| XX | 13-SEP-2002 (first entry) |
| DE | |
| XX | Human prostate expression marker cDNA 21832. |
| XX | |
| XX | Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; |
| KW | pharmacogenomic marker; gene; ss. |
| XX | |
| XX | Homo sapiens. |
| XX | |
| PN | WO200160860-A2. |
| XX | |
| XX | 23-AUG-2001. |
| XX | |
| PF | 20-FEB-2001; 2001WO-US005171. |
| XX | |
| XX | 17-FEB-2000; 2000US-0183319P. |
| PR | 16-MAR-2000; 2000US-0189862P. |
| PR | 25-MAY-2000; 2000US-0207454P. |
| PR | 09-JUN-2000; 2000US-0211314P. |
| PR | 18-JUL-2000; 2000US-0219007P. |
| PR | 13-DEC-2000; 2000US-0255281P. |
| XX | |
| PA | (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC. |
| XX | |
| PI | Schlegel R, Endege WO, Monahan JE; |
| XX | |
| XX | WPI; 2001-662795/76. |
| XX | |
| PT | Novel isolated nucleic acid molecule associated with cancerous state of |
| PT | prostate cells and correlating with presence of prostate cancer, useful |
| PT | for detecting presence of prostate cancer, stage of prostate cancer. |
| XX | |
| PS | Claim 1; Page 3700-3701; 117500p; English. |

| Accession | Sequence | Position |
|-----------|---------------------------------------------------------------------|----------|
| Qy | 101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTTCAATAAGAAAGTAAACCAAGTGTTG | 159 |
| Db | 2095 TATATTATTAACTGTGAAGTATTTTGGCTGATTATAGAAATATAAGATTAACAACAGGGTTA | 2154 |
| Qy | 160 CAACATGCTGTTTAAATAACCCGCAAACTTCAATCACTATA | 200 |
| Db | 2155 CAGTTGTGCAAAACAAAAACAACATACCTTAAATTTACAATTA | 2195 |

RESULT 7
ABV24590
ID ABV24590 standard; cDNA; 2224 BP.
XX
XX AC
XX ABV24590;
XX
XX
XX
DT 16-SEP-2002 (first entry)
XX
XX
XX Human prostate expression marker cDNA 24581.
XX
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200160860-A2.
FN
XX
XX
XX 23-AUG-2001.
PD
XX
XX 20-FEB-2001; 2001WO-US005171.
PF
XX
XX
XX 17-FEB-2000; 2000US-0183319P.
PR
XX 16-MAR-2000; 2000US-0189862P.
PR
XX 25-MAY-2000; 2000US-0207454P.
PR
XX 09-JUN-2000; 2000US-0211314P.
PR
XX 18-JUL-2000; 2000US-0219007P.
PR
XX 13-DEC-2000; 2000US-0255281P.
XX
(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX
XX Schlegel R, Endege WO, Monahan JE;
PI
XX
XX WPI; 2001-662795/76.
DR
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 4665-4666; 11750pp; English.
PS

The invention relates to an isolated nucleic acid molecule. (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;

| | | | | |
|-----------------------|------------------------------------------------------------------|----------------|-----------|--------------|
| Query Match | 12.9% | Score 38.6; | DB 5; | Length 2224; |
| Best Local Similarity | 56.5%; | Pred. No. 1.3; | | |
| Matches 91; | Conservative 0; | Mismatches 69; | Indels 1; | Gaps 1; |
| Qy 41 | TTGAAATATTCTTCATTGTGAAATAAAGACGTAACTCCCAAATTTGCTAACTAAGGGGAG | 100 | | |
| Db 2035 | TTGAAATATTATTTAAAGACTTTAAATTCCTGCTGGTTTGCAAATTCGTATATAACATTATATC | 2094 | | |

QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATTAAGAAGTAACAGAGTGTG 159
 DB 2095 TATATTATTAAGTGAAGTATTTTTCCTGATTATGAATATAAAGTAAACAGGGTTA 2154
 QY 160 CAACATGCTGTTTAAATACCCGACAACTTCAATCACTATA 200
 DB 2155 CAGTTGTGCAAAACAAACAAATACCTTAATTAACAATA 2195

RESULT 8
 ID ABV28008 standard; cDNA; 2224 BP.
 XX AC ABV28008;
 XX DT 16-SEP-2002 (first entry)
 XX DE Human prostate expression marker cDNA 27999.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200160860-A2.
 XX PD 23-AUG-2001.
 XX PF 20-FEB-2001; 2001WO-US005171.
 XX PR 17-FEB-2000; 2000US-0183319P.
 XX PR 16-MAR-2000; 2000US-0189862P.
 XX PR 25-MAY-2000; 2000US-0207454P.
 XX PR 09-JUN-2000; 2000US-0211314P.
 XX PR 18-JUL-2000; 2000US-0219007P.
 XX PR 13-DEC-2000; 2000US-0255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 Schlegel R, Endege WO, Monahan JE;
 WPI; 2001-662795/76.
 Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.
 Claim 1; Page 5779; 11750pp; English.
 The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the efficacy of a test compound to inhibit prostate cancer in a patient; (c) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (d) assessing the efficacy of a composition for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;
 Query Match 12.9%; Score 38.6; DB 5; Length 2224;
 Best Local Similarity 56.5%; Pred. No. 1.3;
 Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
 QY 41 TTGAAATATTCTTCATTTGAAATAAAGCAGCTAACTCCCAATTTGCTTAACCTAGGGAG 100
 DB 2035 TTGAATATTATTTAAAGCTTAAATTCCTGCTGGTTTGCATTTGATATAATATATC 2094
 QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATTAAGAAGTAACAGAGTGTG 159

DB 2095 TATATTATTAAGTGAAGTATTTTTCCTGATTATGAATATAAAGTAAACAGGGTTA 2154
 QY 160 CAACATGCTGTTTAAATACCCGACAACTTCAATCACTATA 200
 DB 2155 CAGTTGTGCAAAACAAACAAATACCTTAATTAACAATA 2195

RESULT 9
 ID ABV23184 standard; cDNA; 2224 BP.
 XX AC ABV23184;
 XX DT 16-SEP-2002 (first entry)
 XX DE Human prostate expression marker cDNA 23175.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200160860-A2.
 XX PD 23-AUG-2001.
 XX PF 20-FEB-2001; 2001WO-US005171.
 XX PR 17-FEB-2000; 2000US-0183319P.
 XX PR 16-MAR-2000; 2000US-0189862P.
 XX PR 25-MAY-2000; 2000US-0207454P.
 XX PR 09-JUN-2000; 2000US-0211314P.
 XX PR 18-JUL-2000; 2000US-0219007P.
 XX PR 13-DEC-2000; 2000US-0255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 Schlegel R, Endege WO, Monahan JE;
 WPI; 2001-662795/76.
 Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.
 Claim 1; Page 4163-4164; 11750pp; English.
 The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the efficacy of a test compound to inhibit prostate cancer in a patient; (c) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (d) assessing the efficacy of a composition for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;
 Query Match 12.9%; Score 38.6; DB 5; Length 2224;
 Best Local Similarity 56.5%; Pred. No. 1.3;
 Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
 QY 41 TTGAAATATTCTTCATTTGAAATAAAGCAGCTAACTCCCAATTTGCTTAACCTAGGGAG 100
 DB 2035 TTGAATATTATTTAAAGCTTAAATTCCTGCTGGTTTGCATTTGATATAATATATC 2094
 QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATTAAGAAGTAACAGAGTGTG 159

| | | | |
|----|------|-----------------------------------------------------------|------|
| Db | 2095 | TATATTAACTGTTAAAGTATTTTTCCTGATTGAAATATAAAGTAACAACAGGGCTTA | 2154 |
| Qy | 160 | CAACATGCTGTTAAATTAACCCGACAAACTTCAATCACTATA | 200 |
| Db | 2155 | CAGTTGTGCAAAACAAACCAAAACATACCTAAATTAACAATA | 2195 |

| | |
|-----------|-------------------------------------------------------------------------|
| RESULT 10 | |
| ABV24513 | |
| ID | ABV24513 standard; cDNA; 2224 BP. |
| XX | |
| AC | ABV24513; |
| XX | |
| DT | 16-SEP-2002 (first entry) |
| XX | |
| DE | Human prostate expression marker cDNA 24504. |
| XX | |
| KW | Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; |
| KW | pharmacogenomic marker; gene; ss. |
| OS | |
| XX | Homo sapiens. |
| XX | |
| PN | WO200160860-A2. |
| XX | |
| XX | 23-AUG-2001. |
| XX | |
| PF | 20-FEB-2001; 2001WO-US005171. |
| XX | |
| PR | 17-FEB-2000; 2000US-0183319P. |
| PR | 16-MAR-2000; 2000US-0189862P. |
| PR | 25-MAY-2000; 2000US-0207454P. |
| PR | 09-JUN-2000; 2000US-0211314P. |
| PR | 18-JUL-2000; 2000US-0219007P. |
| PR | 13-DEC-2000; 2000US-0255281P. |
| XX | |
| PA | (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC. |

| | Query Match | 12.9% | Score 38.6; | DB 5; | Length 2224; |
|----|-----------------------|------------------------------------------------------------------|-----------------|-----------|--------------|
| | Best Local Similarity | 56.5% | Pred. N1.3; | | |
| | Matches 91; | Conservative 0; | No. Matches 69; | Indels 1; | Gaps 1; |
| Qy | 41 | TTGAAATATTCTTCATTGAAATAAACAGCTAACTCCCAAATGTGACCTAACTAGGGGAG | 100 | | |
| Db | 2035 | TTGAAATATTTTAAAGCTTAAAAATTCGCTGGTTTGCAAAATGTATATAACATTATATC | 2094 | | |
| Qy | 101 | TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTTCAAATAAGAAAGTAACCAAGAGTGTGTG | 159 | | |
| Db | 2095 | TATATATTAACTGCTGACTATTTTGGCTGATATGTAATATAAGTAATCAACAGGGTTA | 2154 | | |

Qy 160 CAACATGCTGTAAATAAACCCGACAAAATTCAATCACATATA 200
|| | ||| ||| ||| ||| ||| |||
pB 2155 CAGTTGTGCAAAACA AAAACAATACCTTAATTAACAATA 2195

| | |
|-----------|-------------------------------------------------------------------------|
| RESULT 11 | |
| ABV24601 | |
| ID | ABV24601 standard; cDNA; 2224 BP. |
| XX | |
| XX | |
| AC | ABV24601; |
| XX | |
| XX | |
| DT | 16-SEP-2002 (first entry) |
| XX | |
| DE | Human prostate expression marker cDNA 24592. |
| XX | |
| XX | |
| KW | Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; |
| KW | pharmacogenomic marker; gene; ss. |
| XX | |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | WO200160860-A2. |
| XX | |
| PD | 23-AUG-2001. |
| XX | |
| XX | |
| PF | 20-FEB-2001; 2001WO-US005171. |
| XX | |
| PR | 17-FEB-2000; 2000US-0183319P. |
| PR | 16-MAR-2000; 2000US-0189862P. |
| PR | 25-MAY-2000; 2000US-0207454P. |
| PR | 09-JUN-2000; 2000US-0211314P. |
| PR | 18-JUL-2000; 2000US-0219007P. |
| PR | 13-DEC-2000; 2000US-0255281P. |
| XX | |
| XX | |
| PA | (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC. |
| XX | |
| PI | Schlegel R, Endege WO, Monahan JE; |
| XX | |
| DR | WPI: 2001-662795/76. |

| | Query Match | 12.9% | Score 38.6; | DB 5; | Length 2224; |
|----|-----------------------|------------------------------------------------------------------|----------------|-----------|--------------|
| | Best Local Similarity | 56.5%; | Pred. No. 1.3; | | |
| | Matches 91; | Conservative 0; | Mismatches 69; | Indels 1; | Gaps 1; |
| QY | 41 | TTGAAATATTCTTCATTTGAAATAAACAGCTAACTCCCAAATTGTACTTACTAGGGAG | 100 | | |
| DB | 2035 | TTGAATATTATTTTAAAGCTTAAAAATCTGCTGGTTTCCAAATTTGTATATACATTATATC | 2094 | | |
| QY | 101 | TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTTCAAATAAGAAAGTAAACACAGAGTGTTG | 159 | | |
| DB | 2095 | TATATATTAACTCTGAAGTATTTTTCCTGATATGAATATGAATATAAAGTAAACAAACGGGTTA | 2154 | | |

Qy 160 CAACATGCTGTTAAATACCCGACAACTTCAATCACTATA 200
 |||
 Ds 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2195
 |||

RESULT 12
 ABV29023
 ID ABV29023 standard; cDNA; 2224 BP.
 XX AC
 XX ABV29023;
 XX DT
 XX 16-SEP-2002 (first entry)
 XX DE Human prostate expression marker cDNA 29014.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX FN WO200160860-A2.
 XX PD 23-AUG-2001.
 XX PE 20-FEB-2001; 2001WO-US005171.
 XX PR 17-FEB-2000; 2000US-0183319P.
 XX PR 16-MAR-2000; 2000US-0189862P.
 XX PR 25-MAY-2000; 2000US-0207454P.
 XX PR 09-JUN-2000; 2000US-0211314P.
 XX PR 18-JUL-2000; 2000US-0219007P.
 XX PR 13-DEC-2000; 2000US-0255281P.
 XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX PI Schlegel R, Endege WO, Monahan JE;
 XX DR WPI; 2001-662795/76.
 XX PT Novel isolated nucleic acid molecule associated with cancerous state of
 XX PT prostate cells and correlating with presence of prostate cancer, useful
 XX PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX PS Claim 1; Page 6143; 11750pp; English.
 XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
 XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 XX CC specification or its complement. (I) is useful for: (a) assessing whether
 XX CC a patient is afflicted with prostate cancer; (b) monitoring the
 XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
 XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 XX CC determining whether prostate cancer has metastasized in a patient; (h)
 XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
 XX CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX SQ Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;

Query Match 12.9%; Score 38.6; DB 5; Length 2224;
 Best Local Similarity 56.5%; Pred. No. 1.3;
 Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
 Qy 41 TTGAATATTTCTTCAATTTGAATTAACAGCTAACTCCCAATTTGCTAACTACTAGGGAG 100
 |||
 Ds 2035 TTGAATATTTTAAAGCTTAAATTTCTGCTGGTTGCAAAATTTGTATATAATATATATC 2094
 |||

Qy 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATTAAGTAAGTAACAGAGTGTG 159
 |||
 Ds 2095 TATATTAATTAAGTAAATTTTGGCTGATTAATGAATTAAGTAAGTAACAGAGGTTA 2154
 |||

Qy 160 CAACATGCTGTTAAATACCCGACAACTTCAATCACTATA 200
 |||

Db 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2195
 |||

RESULT 13
 ABV28014
 ID ABV28014 standard; cDNA; 2224 BP.
 XX AC
 XX ABV28014;
 XX DT
 XX 16-SEP-2002 (first entry)
 XX DE Human prostate expression marker cDNA 28005.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX FN WO200160860-A2.
 XX PD 23-AUG-2001.
 XX PE 20-FEB-2001; 2001WO-US005171.
 XX PR 17-FEB-2000; 2000US-0183319P.
 XX PR 16-MAR-2000; 2000US-0189862P.
 XX PR 25-MAY-2000; 2000US-0207454P.
 XX PR 09-JUN-2000; 2000US-0211314P.
 XX PR 18-JUL-2000; 2000US-0219007P.
 XX PR 13-DEC-2000; 2000US-0255281P.
 XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX PI Schlegel R, Endege WO, Monahan JE;
 XX DR WPI; 2001-662795/76.
 XX PT Novel isolated nucleic acid molecule associated with cancerous state of
 XX PT prostate cells and correlating with presence of prostate cancer, useful
 XX PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX PS Claim 1; Page 5782; 11750pp; English.
 XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
 XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 XX CC specification or its complement. (I) is useful for: (a) assessing whether
 XX CC a patient is afflicted with prostate cancer; (b) monitoring the
 XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
 XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 XX CC determining whether prostate cancer has metastasized in a patient; (h)
 XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
 XX CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX SQ Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;

Query Match 12.9%; Score 38.6; DB 5; Length 2224;
 Best Local Similarity 56.5%; Pred. No. 1.3;
 Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
 Qy 41 TTGAATATTTCTTCAATTTGAATTAACAGCTAACTCCCAATTTGCTAACTACTAGGGAG 100
 |||
 Ds 2035 TTGAATATTTTAAAGCTTAAATTTCTGCTGGTTGCAAAATTTGTATATAATATATATC 2094
 |||

Qy 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATTAAGTAAGTAACAGAGTGTG 159
 |||
 Ds 2095 TATATTAATTAAGTAAATTTTGGCTGATTAATGAATTAAGTAAGTAACAGAGGTTA 2154
 |||

Qy 160 CAACATGCTGTTAAATACCCGACAACTTCAATCACTATA 200
 |||

Db 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2195

RESULT 14

ABV22169
ID ABV22169 standard; cDNA; 2224 BP.

XX
AC ABV22169;

XX
DT 13-SEP-2002 (first entry)

XX
DE Human prostate expression marker cDNA 22160.

XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX
KW pharmacogenomic marker; gene; ss.

XX
OS Homo sapiens.

XX
PN W0200160860-A2.

XX
PD 23-AUG-2001.

XX
PF 20-FEB-2001; 2001WO-US005171.

XX
PR 17-FEB-2000; 2000US-0183319P.

XX
PR 16-MAR-2000; 2000US-0189862P.

XX
PR 25-MAY-2000; 2000US-0207454P.

XX
PR 09-JUN-2000; 2000US-0211314P.

XX
PR 18-JUL-2000; 2000US-0219007P.

XX
PR 13-DEC-2000; 2000US-0255281P.

XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX
PI Schlegel R, Endege WO, Monahan JE;

XX
DR WPI; 2001-662795/76.

XX
PT Novel isolated nucleic acid molecule associated with cancerous state of

XX
PT prostate cells and correlating with presence of prostate cancer, useful

XX
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX
PS Claim 1; Page 3811; 11750pp; English.

XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX
CC specification or its complement. (I) is useful for: (a) assessing whether
XX
CC a patient is afflicted with prostate cancer; (b) monitoring the
XX
CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX
CC determining whether prostate cancer has metastasized in a patient; (h)
XX
CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX
SQ Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;

Query Match 12.9%; Score 38.6; DB 5; Length 2224;
Best Local Similarity 56.5%; Pred. No. 1.3;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

Qy 41 TTGAATATTTTCAATTTGAAATAAACAGCTAACTCCCAATTTGCTTAACCTAGGGGAG 100
|||||
2035 TTGAATATTTTCAATTTGAAATTTCTGCTGGTTTGCATTTGTATATAACATTATATC 2094

Db 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATTAAGTAAGTAACAGAGTGTG 159
|||||

Qy 2095 TATATTTAACTGTAAGTATTTTGGCTGATTAAGTAAGTAACAGAGGTTA 2154
|||||

Qy 160 CAACATGCTGTTAAATACCCGACAACTTCAATCACTATA 200
|||||

Db 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2195
|||||

RESULT 15

ABV22175
ID ABV22175 standard; cDNA; 2224 BP.

XX
AC ABV22175;

XX
DT 13-SEP-2002 (first entry)

XX
DE Human prostate expression marker cDNA 22166.

XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX
KW pharmacogenomic marker; gene; ss.

XX
OS Homo sapiens.

XX
PN W0200160860-A2.

XX
PD 23-AUG-2001.

XX
PF 20-FEB-2001; 2001WO-US005171.

XX
PR 17-FEB-2000; 2000US-0183319P.

XX
PR 16-MAR-2000; 2000US-0189862P.

XX
PR 25-MAY-2000; 2000US-0207454P.

XX
PR 09-JUN-2000; 2000US-0211314P.

XX
PR 18-JUL-2000; 2000US-0219007P.

XX
PR 13-DEC-2000; 2000US-0255281P.

XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX
PI Schlegel R, Endege WO, Monahan JE;

XX
DR WPI; 2001-662795/76.

XX
PT Novel isolated nucleic acid molecule associated with cancerous state of

XX
PT prostate cells and correlating with presence of prostate cancer, useful

XX
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX
PS Claim 1; Page 3812-3813; 11750pp; English.

XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX
CC specification or its complement. (I) is useful for: (a) assessing whether
XX
CC a patient is afflicted with prostate cancer; (b) monitoring the
XX
CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX
CC determining whether prostate cancer has metastasized in a patient; (h)
XX
CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX
SQ Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;

Query Match 12.9%; Score 38.6; DB 5; Length 2224;
Best Local Similarity 56.5%; Pred. No. 1.3;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

Qy 41 TTGAATATTTTCAATTTGAAATAAACAGCTAACTCCCAATTTGCTTAACCTAGGGGAG 100
|||||
2035 TTGAATATTTTCAATTTGAAATTTCTGCTGGTTTGCATTTGTATATAACATTATATC 2094

Db 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATTAAGTAAGTAACAGAGTGTG 159
|||||

Qy 2095 TATATTTAACTGTAAGTATTTTGGCTGATTAAGTAAGTAACAGAGGTTA 2154
|||||

Qy 160 CAACATGCTGTTAAATACCCGACAACTTCAATCACTATA 200
|||||

Db 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2195
|||||

```
RESULT 16
ABV25057
ID ABV25057 standard; cDNA; 2224 BP.
XX AC
XX ABV25057;
XX DT
XX 16-SEP-2002 (first entry)
XX DE
XX Human prostate expression marker cDNA 25048.
XX KW
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX KW
XX OS
XX Homo sapiens.
XX PN
XX WO200160860-A2.
XX PD
XX 23-AUG-2001.
XX PF
XX 20-FEB-2001; 2001WO-US005171.
XX PR
XX 17-FEB-2000; 2000US-0183319P.
XX PR
XX 16-MAR-2000; 2000US-0189862P.
XX PR
XX 25-MAY-2000; 2000US-0207454P.
XX PR
XX 09-JUN-2000; 2000US-0211314P.
XX PR
XX 18-JUL-2000; 2000US-0219007P.
XX PR
XX 13-DEC-2000; 2000US-0255281P.
XX PA
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI
XX Schlegel R, Endege WO, Monahan JE;
XX DR
XX WPI; 2001-662795/76.
XX PT
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS
XX Claim 1; Page 4864; 11750pp; English.
XX CC
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)
XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX CC
XX SQ
XX Sequence 2224 BP; 661 A; 410 C; 419 G; 723 T; 0 U; 11 Other;
XX CC
XX Query Match 12.9%; Score 38.6; DB 5; Length 2224;
XX CC
XX Best Local Similarity 56.5%; Pred. No. 1.3;
XX CC
XX Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
XX CC
XX 41 TTGAATATTCTTCATTGTAATAAAGCAGCTAACTCCCAATTTGCTACTACTAGGGAG 100
XX DB
XX 2035 TTGAATATTCTTAAAGCTTAAATTTCTGCTGTTTGAATTTGTAATACTATATATC 2094
XX CC
XX 101 TTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAAATGAAGTAACAGAGTGTG 159
XX DB
XX 2095 TATATTATTAAGTAAATTTTTCCTGCTGATTGAAATATAAAGTAACACAGGGTTA 2154
XX CC
XX 160 CAACATGCTGTTAAATACCGCAGCAAACTTCAATCACTATA 200
XX DB
XX 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2195
XX CC
XX Query Match 12.4%; Score 37.2; DB 6; Length 14920;
XX CC
XX Best Local Similarity 49.0%; Pred. No. 5.5;
XX CC
XX Matches 99; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
XX CC
XX 40 CTTGAATATTCTTCATTGCAATATAACAGCTAACTCCCAATTTGCTACTAGGGGA 99
XX DB
XX 12920 CATAAATAACAATAATTTTATATAAAACATAAACTAACTAAATATAACGAACGT 12861
XX CC
XX 100 GTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAAATGAAGTAACAGAGTGTG 159
XX DB
XX 12860 ATATACGAATAAATAAAATTTAAATAATTAACACTCATTAAAAACATATCTATTAAATTC 12801
XX CC
XX Sequence 14920 BP; 4132 A; 171 C; 3248 G; 7369 T; 0 U; 0 Other;
XX CC
XX Query Match 12.4%; Score 37.2; DB 6; Length 14920;
XX CC
XX Best Local Similarity 49.0%; Pred. No. 5.5;
XX CC
XX Matches 99; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
XX CC
XX 40 CTTGAATATTCTTCATTGCAATATAACAGCTAACTCCCAATTTGCTACTAGGGGA 99
XX DB
XX 12920 CATAAATAACAATAATTTTATATAAAACATAAACTAACTAAATATAACGAACGT 12861
XX CC
XX 100 GTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAAATGAAGTAACAGAGTGTG 159
XX DB
XX 12860 ATATACGAATAAATAAAATTTAAATAATTAACACTCATTAAAAACATATCTATTAAATTC 12801
XX CC
XX The invention relates to a nucleic acid (I) comprising a sequence at
XX CC least 18 bases in length of a segment of chemically pretreated DNA (II)
XX CC of genes associated with development selected from 87 genes listed in the
XX CC specification such as ACCFN, ADFN, or AFD1 and comprising one of 350
XX CC sequences (ABN79984-ABN80333) or their complements. The invention is
XX CC useful for the diagnosis or therapy of diseases associated with
XX CC development genes, in particular disease related to homeobox containing
XX CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
XX CC associated with congenital heart disease, epilepsy, diseases related to
XX CC histone deacetylation, Currarino syndrome, diseases related with the
XX CC development of the brain and limb girdle muscular dystrophy and dwarfism.
XX CC Oligomers specific to each of the genes are useful for detecting the
XX CC methylation state of all CpG dinucleotides within the 350 sequences or
XX CC (II) and their complementary sequences, as primer oligonucleotides for
XX CC the amplification of the 350 sequences, (II) and/or their complements and
XX CC as oligomer probes for detecting the cytosine methylation state and/or
XX CC single nucleotide polymorphisms (SNPs). Note: The sequence data for this
XX CC patent did not form part of the printed specification but is based on
XX CC sequence information supplied to Derwent by the European Patent Office
XX CC
XX SQ
XX Sequence 14920 BP; 4132 A; 171 C; 3248 G; 7369 T; 0 U; 0 Other;
XX CC
XX Query Match 12.4%; Score 37.2; DB 6; Length 14920;
XX CC
XX Best Local Similarity 49.0%; Pred. No. 5.5;
XX CC
XX Matches 99; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
XX CC
XX 40 CTTGAATATTCTTCATTGCAATATAACAGCTAACTCCCAATTTGCTACTAGGGGA 99
XX DB
XX 12920 CATAAATAACAATAATTTTATATAAAACATAAACTAACTAAATATAACGAACGT 12861
XX CC
XX 100 GTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAAATGAAGTAACAGAGTGTG 159
XX DB
XX 12860 ATATACGAATAAATAAAATTTAAATAATTAACACTCATTAAAAACATATCTATTAAATTC 12801
```

160 CAACATGCTGTTAAATAAACCCGACAACTTCAATCACTATAGCTGTAGTAGAGTGCAATC 219
 12800 CAATATTTTCTAAAAAATAAATAATCTCTTAACCAAAAAATTTCACTAACTCAATACATCA 12741

220 TGCAAGATCCAGAGTACCA 241
 12740 CCAATAATATCTCAAAACCA 12719

RESULT 18
 ADW53569/C
 ID ADW53569 standard; cDNA; 6857 BP.
 XX
 AC ADW53569;
 XX
 DT 24-MAR-2005 (first entry)
 XX
 DE Human 98P4B6 cDNA variant 30 SeqID 60.
 XX
 KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
 KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
 KW cytosstatic; gene; chromosome 7q21.
 XX
 OS Homo sapiens.
 XX
 PN WO2003087306-A2.
 XX
 PD 23-OCT-2003.
 XX
 PF 04-APR-2003; 2003WO-US010462.
 XX
 PR 05-APR-2002; 2002US-0370387P.
 PR 06-JUN-2002; 2002US-00165044.
 PR 20-DEC-2002; 2002US-0435480P.
 XX
 PA (AGEN-) AGENYSYS INC.
 XX
 PI Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
 XX
 DR WPI; 2003-903158/82.
 DR P-PSDB; ADW53570.
 XX
 PT A composition comprising 98P4B6 proteins and nucleic acid molecules for
 PT detecting, preventing, prognosing and/or treating cancers that express
 PT 98P4B6 (e.g. pancreatic, prostatic or lung cancer).
 XX
 PS Example 2; SEQ ID NO 60; 616pp; English.
 XX
 CC This invention relates to a novel human gene 98P4B6 (also known as STEAP-
 CC 2, six transmembrane epithelial antigen of prostate-2) located on
 CC chromosome 7q21 and the encoded protein and variants derived thereof.
 CC Specifically, it refers to the expression pattern of this gene in adult
 CC normal tissues and its aberrant over-expression in various cancers
 CC including breast, colon, lung, kidney and prostate. The present invention
 CC describes compositions and methods useful for detecting, preventing,
 CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the
 CC polynucleotide can be used for gene therapy purposes, for monitoring
 CC genetic abnormalities and for generating transgenic or 'knockout',
 CC animals, which can be useful for the development and screening of
 CC therapeutically useful reagents. The encoded proteins may also be used in
 CC generating antibodies and vaccines, as well as in identifying ligands and
 CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
 CC these compositions exhibit cytostatic activities. This polynucleotide
 CC sequence is a human 98P4B6 cDNA variant of the invention.
 XX
 SQ Sequence 6857 BP; 2019 A; 1318 C; 1210 G; 2310 T; 0 U; 0 Other;
 Query Match 12.3%; Score 37; DB 11; Length 6857;
 Best Local Similarity 55.9%; Pred. No. 5.1;
 Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

41 TTGAAATATTTCTTGAATAAATAACAGCTAACTCCAAATTTGTTACTTAAGTGGGAG 100

2234 TTGAATATTTTAAAAAGCTTAAATTTCTGCTGTTTGCAAATTTGTTTAAACATTATATC 2175
 101 TTTATCATCAA-GAAATGTTTAAATGCTTAAATTTTCAAATAAGAGTAACACAGAGTGTG 159
 2174 TATATTTAACTGTAAGTATTTTTCCTGATTAATGAATATAAAGTAACAACAGGGTTA 2115

160 CAACATGCTGTTAAATAAACCCGACAACTTCAATCACTATA 200
 2114 CAGTTGCAAAACAAACAAACATACCTTAATTAACAATA 2074

RESULT 19
 ADW53577/C
 ID ADW53577 standard; cDNA; 6857 BP.
 XX
 AC ADW53577;
 XX
 DT 24-MAR-2005 (first entry)
 XX
 DE Human 98P4B6 cDNA variant 34 SeqID 68.
 XX
 KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
 KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
 KW cytosstatic; gene; chromosome 7q21.
 XX
 OS Homo sapiens.
 XX
 PN WO2003087306-A2.
 XX
 PD 23-OCT-2003.
 XX
 PF 04-APR-2003; 2003WO-US010462.
 XX
 PR 05-APR-2002; 2002US-0370387P.
 PR 06-JUN-2002; 2002US-00165044.
 PR 20-DEC-2002; 2002US-0435480P.
 XX
 PA (AGEN-) AGENYSYS INC.
 XX
 PI Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
 XX
 DR WPI; 2003-903158/82.
 DR P-PSDB; ADW53578.
 XX
 PT A composition comprising 98P4B6 proteins and nucleic acid molecules for
 PT detecting, preventing, prognosing and/or treating cancers that express
 PT 98P4B6 (e.g. pancreatic, prostatic or lung cancer).
 XX
 PS Example 2; SEQ ID NO 68; 616pp; English.
 XX
 CC This invention relates to a novel human gene 98P4B6 (also known as STEAP-
 CC 2, six transmembrane epithelial antigen of prostate-2) located on
 CC chromosome 7q21 and the encoded protein and variants derived thereof.
 CC Specifically, it refers to the expression pattern of this gene in adult
 CC normal tissues and its aberrant over-expression in various cancers
 CC including breast, colon, lung, kidney and prostate. The present invention
 CC describes compositions and methods useful for detecting, preventing,
 CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the
 CC polynucleotide can be used for gene therapy purposes, for monitoring
 CC genetic abnormalities and for generating transgenic or 'knockout',
 CC animals, which can be useful for the development and screening of
 CC therapeutically useful reagents. The encoded proteins may also be used in
 CC generating antibodies and vaccines, as well as in identifying ligands and
 CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
 CC these compositions exhibit cytostatic activities. This polynucleotide
 CC sequence is a human 98P4B6 cDNA variant of the invention.
 XX
 SQ Sequence 6857 BP; 2018 A; 1317 C; 1210 G; 2312 T; 0 U; 0 Other;
 Query Match 12.3%; Score 37; DB 11; Length 6857;
 Best Local Similarity 55.9%; Pred. No. 5.1;
 Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

QY 41 TTGAATATCTTCAATTTTGAATAAAGCAGTAACCTCCAAATTTGACCTAATAGGGGAG 100
 DB 2234 TTGAATATTTTAAAGCTTAAATTTCTGCTGTTTGCATAATTTGTTTAAACATTATATC 2175
 QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAGTAACAGAGTGTG 159
 DB 2174 TATATTTAATCTGTAAGTATTTTTCCTGCTGATATGAAATATAAAGTAACACAGGGTTA 2115
 QY 160 CAACATGCTGTTAAATAACCCGACAACTTCAATCACTATA 200
 DB 2114 CAGTTGTGCAAAACAAACACATCACTTAATTAACAATA 2074

RESULT 20

ADW53581/c
 ID ADW53581 standard; cDNA; 6857 BP.

XX AC ADW53581;

XX DT 24-MAR-2005 (first entry)

XX DE Human 98P4B6 cDNA variant 36 SeqID 72.

XX KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
 KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
 XX KW cytosolic; gene; chromosome 7q21.

XX OS Homo sapiens.

XX PN WO2003087306:A2.

XX PD 23-OCT-2003.

XX PF 04-APR-2003; 2003WO-US010462.

XX PR 05-APR-2002; 2002US-0370387P.

XX PR 06-JUN-2002; 2002US-00165044.

XX PR 20-DEC-2002; 2002US-0435480P.

XX PA (AGEN-) AGENSYS INC.

XX PI Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;

XX DR WPI; 2003-903158/82.

XX DR P-PSDB; ADW53582.

XX PT A composition comprising 98P4B6 proteins and nucleic acid molecules for
 PT detecting, preventing, prognosing and/or treating cancers that express
 XX 98P4B6 (e.g. pancreatic, prostatic or lung cancer).

XX PS Example 2; SEQ ID NO 72; 616pp; English.

XX PS This invention relates to a novel human gene 98P4B6 (also known as STEAP-
 CC 2, six transmembrane epithelial antigen of prostate-2) located on
 CC chromosome 7q21 and the encoded protein and variants derived thereof.
 CC Specifically, it refers to the expression pattern of this gene in adult
 CC normal tissues and its aberrant over-expression in various cancers
 CC including breast, colon, lung, kidney and prostate. The present invention
 CC describes compositions and methods useful for detecting, preventing,
 CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the
 CC polynucleotide can be used for gene therapy purposes, for monitoring
 CC genetic abnormalities and for generating transgenic or 'knockout'
 CC animals, which can be useful for the development and screening of
 CC therapeutically useful reagents. The encoded proteins may also be used in
 CC generating antibodies and vaccines, as well as in identifying ligands and
 CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
 CC these compositions exhibit cytostatic activities. This polynucleotide
 CC sequence is a human 98P4B6 cDNA variant of the invention.

XX PS Sequence 6857 BP; 2018 A; 1317 C; 1211 G; 2311 T; 0 U; 0 Other;

XX PS Query Match 12.3%; Score 37; DB 11; Length 6857;

Best Local Similarity 55.9%; Pred. No. 5.1;
 Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
 QY 41 TTGAATATCTTCAATTTTGAATAAAGCAGTAACCTCCAAATTTGACCTAATAGGGGAG 100
 DB 2234 TTGAATATTTTAAAGCTTAAATTTCTGCTGTTTGCATAATTTGTTTAAACATTATATC 2175
 QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAGTAACAGAGTGTG 159
 DB 2174 TATATTTAATCTGTAAGTATTTTTCCTGCTGATATGAAATATAAAGTAACACAGGGTTA 2115
 QY 160 CAACATGCTGTTAAATAACCCGACAACTTCAATCACTATA 200
 DB 2114 CAGTTGTGCAAAACAAACACATCACTTAATTAACAATA 2074

RESULT 21

ADW53696/c

ID ADW53696 standard; cDNA; 6857 BP.

XX AC ADW53696;

XX DT 24-MAR-2005 (first entry)

XX DE Partial human 98P4B6 cDNA variant 8 SeqID 187.

XX KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
 KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
 XX KW cytosolic.

XX OS Homo sapiens.

XX PN WO2003087306-A2.

XX PD 23-OCT-2003.

XX PF 04-APR-2003; 2003WO-US010462.

XX PR 05-APR-2002; 2002US-0370387P.

XX PR 06-JUN-2002; 2002US-00165044.

XX PR 20-DEC-2002; 2002US-0435480P.

XX PA (AGEN-) AGENSYS INC.

XX PI Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;

XX DR WPI; 2003-903158/82.

XX DR P-PSDB; ADW53699; ADW53701.

XX PT A composition comprising 98P4B6 proteins and nucleic acid molecules for
 PT detecting, preventing, prognosing and/or treating cancers that express
 XX 98P4B6 (e.g. pancreatic, prostatic or lung cancer).

XX PS Example 51; SEQ ID NO 187; 616pp; English.

XX PS This invention relates to a novel human gene 98P4B6 (also known as STEAP-
 CC 2, six transmembrane epithelial antigen of prostate-2) located on
 CC chromosome 7q21 and the encoded protein and variants derived thereof.
 CC Specifically, it refers to the expression pattern of this gene in adult
 CC normal tissues and its aberrant over-expression in various cancers
 CC including breast, colon, lung, kidney and prostate. The present invention
 CC describes compositions and methods useful for detecting, preventing,
 CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the
 CC polynucleotide can be used for gene therapy purposes, for monitoring
 CC genetic abnormalities and for generating transgenic or 'knockout'
 CC animals, which can be useful for the development and screening of
 CC therapeutically useful reagents. The encoded proteins may also be used in
 CC generating antibodies and vaccines, as well as in identifying ligands and
 CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
 CC these compositions exhibit cytostatic activities. This polynucleotide
 CC sequence is a partial human 98P4B6 cDNA variant of the invention.

XX PS Sequence 6857 BP; 2018 A; 1318 C; 1210 G; 2311 T; 0 U; 0 Other;

CC animals, which can be useful for the development and screening of
CC therapeutically useful reagents. The encoded proteins may also be used in
CC generating antibodies and vaccines, as well as in identifying ligands and
CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
CC these compositions exhibit cytostatic activities. This polynucleotide
CC sequence is a human 98P4B6 cDNA variant of the invention.

XX Sequence 6857 BP; 2017 A; 1319 C; 1210 G; 2311 T; 0 U; 0 Other;

SQ Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
QY 41 TTGAATATTCTTCATTTGAAATAAACAGCTAACTCCCAATTTGCTAACTAGGGAG 100
DB 2234 TTGAATATTCTTCATTTGAAATAAACAGCTAACTCCCAATTTGCTAACTAGGGAG 100
QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTCAATAAGAGTAACAGAGTGTG 159
DB 2174 TATATTATTAACTGTAAGTATTTTTCCTGCTGATTATGAATATATAAGTAACAACAGGGTTA 2115
QY 160 CAACATGCTGTGTTAAATAACCGCAACAACTTCAATCACTATA 200
DB 2114 CAGTTGTGCAAAACAAACAACTACCTAAATTACAATA 2074

RESULT 26

ADW53575/c
ID ADW53575 standard; cDNA; 6857 BP.

XX AC ADW53575;

XX DT 24-MAR-2005 (first entry)

XX DE Human 98P4B6 cDNA variant 33 SeqID 66.

XX KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
XX KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
XX KW cytostatic; gene; chromosome 7q21.

XX OS Homo sapiens.

XX PN WO2003087306-A2.

XX PD 23-OCT-2003.

XX PF 04-APR-2003; 2003WO-US010462.

XX PR 05-APR-2002; 2002US-0370387P.

XX PR 06-JUN-2002; 2002US-00165044.

XX PR 20-DEC-2002; 2002US-0435480P.

XX PA (AGEN-) AGENYSYS INC.

XX PI Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;

XX WPI; 2003-903158/82.

XX P-PSDB; ADW53576.

XX A composition comprising 98P4B6 proteins and nucleic acid molecules for
XX detecting, preventing, prognosing and/or treating cancers that express
XX 98P4B6 (e.g. pancreatic, prostatic or lung cancer).

XX Example 2; SEQ ID NO 66; 616pp; English.

XX This invention relates to a novel human gene 98P4B6 (also known as STEAP-
XX 2, six transmembrane epithelial antigen of prostate-2) located on
XX chromosome 7q21 and the encoded protein and variants derived thereof.
XX Specifically, it refers to the expression pattern of this gene in adult
XX normal tissues and its aberrant over-expression in various cancers
XX including breast, colon, lung, kidney and prostate. The present invention
XX describes compositions and methods useful for detecting, preventing,
XX prognosing and/or treating cancers that express 98P4B6. Furthermore, the

CC polynucleotide can be used for gene therapy purposes, for monitoring
CC genetic abnormalities and for generating transgenic or 'knockout'
CC animals, which can be useful for the development and screening of
CC therapeutically useful reagents. The encoded proteins may also be used in
CC generating antibodies and vaccines, as well as in identifying ligands and
CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
CC these compositions exhibit cytostatic activities. This polynucleotide
CC sequence is a human 98P4B6 cDNA variant of the invention.

XX Sequence 6857 BP; 2017 A; 1318 C; 1211 G; 2311 T; 0 U; 0 Other;

SQ Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
QY 41 TTGAATATTCTTCATTTGAAATAAACAGCTAACTCCCAATTTGCTAACTAGGGAG 100
DB 2234 TTGAATATTCTTCATTTGAAATAAACAGCTAACTCCCAATTTGCTAACTAGGGAG 100
QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTCAATAAGAGTAACAGAGTGTG 159
DB 2174 TATATTATTAACTGTAAGTATTTTTCCTGCTGATTATGAATATATAAGTAACAACAGGGTTA 2115
QY 160 CAACATGCTGTGTTAAATAACCGCAACAACTTCAATCACTATA 200
DB 2114 CAGTTGTGCAAAACAAACAACTACCTAAATTACAATA 2074

RESULT 27

ADW53571/c
ID ADW53571 standard; cDNA; 6857 BP.

XX AC ADW53571;

XX DT 24-MAR-2005 (first entry)

XX DE Human 98P4B6 cDNA variant 31 SeqID 62.

XX KW ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
XX KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
XX KW cytostatic; gene; chromosome 7q21.

XX OS Homo sapiens.

XX PN WO2003087306-A2.

XX PD 23-OCT-2003.

XX PF 04-APR-2003; 2003WO-US010462.

XX PR 05-APR-2002; 2002US-0370387P.

XX PR 06-JUN-2002; 2002US-00165044.

XX PR 20-DEC-2002; 2002US-0435480P.

XX PA (AGEN-) AGENYSYS INC.

XX PI Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;

XX WPI; 2003-903158/82.

XX P-PSDB; ADW53572.

XX A composition comprising 98P4B6 proteins and nucleic acid molecules for
XX detecting, preventing, prognosing and/or treating cancers that express
XX 98P4B6 (e.g. pancreatic, prostatic or lung cancer).

XX Example 2; SEQ ID NO 62; 616pp; English.

XX This invention relates to a novel human gene 98P4B6 (also known as STEAP-
XX 2, six transmembrane epithelial antigen of prostate-2) located on
XX chromosome 7q21 and the encoded protein and variants derived thereof.
XX Specifically, it refers to the expression pattern of this gene in adult
XX normal tissues and its aberrant over-expression in various cancers
XX including breast, colon, lung, kidney and prostate. The present invention

CC describes compositions and methods useful for detecting, preventing, CC
 CC prognosing and/ or treating cancers that express 98P4B6. Furthermore, the CC
 CC polynucleotide can be used for gene therapy purposes, for monitoring, CC
 CC genetic abnormalities and for generating transgenic or 'knockout' CC
 CC animals, which can be useful for the development and screening of CC
 CC therapeutically useful reagents. The encoded proteins may also be used in CC
 CC generating antibodies and vaccines, as well as in identifying ligands and CC
 CC cellular constituents that bind to 98P4B6 gene products. Accordingly, CC
 CC these compositions exhibit cytostatic activities. This polynucleotide CC
 CC sequence is a human 98P4B6 cDNA variant of the invention. CC
 XX
 SQ Sequence 6857 BP; 2017 A; 1318 C; 1211 G; 2311 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
 Best Local Similarity 55.9%; Pred. No. 5.1;
 Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
 QY 41 TTGAAATATCTTCATTTGAAATAAACAGCTAACTCCCAATTTGCTAACTAGGAG 100
 DB 2234 TTGAATATTTTAAAGCTTAAATTTCTGCTGTTTGCATTTGTTTAACTATATC 2175
 QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATTAAGAGTAACAGAGTGTG 159
 DB 2174 TATATTAATTAAGTAAATTTTGGCTGATTAATGAATATAAGTAAACACAGGGTTA 2115
 QY 160 CAACATGCTGTTAAATACCGACAACTTCAATCACTATA 200
 DB 2114 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2074

RESULT 28
 ADW53698/c
 ID ADW53698 standard; cDNA; 6857 BP.
 XX
 AC ADW53698;
 XX
 DT 24-MAR-2005 (first entry)
 XX
 XX Partial human 98P4B6 cDNA variant 8 SeqID 189.
 XX
 XX ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
 KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
 KW cytostatic.
 XX
 OS Homo sapiens.
 XX
 XX WO2003087306-A2.
 XX
 XX 23-OCT-2003.
 XX
 XX 04-APR-2003; 2003WO-US010462.
 XX
 XX 05-APR-2002; 2002US-0370387P.
 XX
 XX 06-JUN-2002; 2002US-00165044.
 XX
 XX 20-DEC-2002; 2002US-0435480P.
 XX
 XX (AGEN-) AGENYSYS INC.
 XX
 XX Challita-Eid PM, Raitano AB, Paris M, Ge W, Jakobovits A;
 XX
 XX WPI; 2003-903158/82.
 XX
 XX P-PSDB; ADW53698, ADW53701.

A composition comprising 98P4B6 proteins and nucleic acid molecules for
 PT detecting, preventing, prognosing and/or treating cancers that express
 PT 98P4B6 (e.g. pancreatic, prostatic or lung cancer).
 XX
 XX Example 51; SEQ ID NO 189; 616pp; English.
 XX
 XX This invention relates to a novel human gene 98P4B6 (also known as STEAP-
 CC 2, six transmembrane epithelial antigen of prostate-2) located on
 CC chromosome 7q21 and the encoded protein and variants derived thereof.
 CC Specifically, it refers to the expression pattern of this gene in adult

CC normal tissues and its aberrant over-expression in various cancers
 CC including breast, colon, lung, kidney and prostate. The present invention
 CC describes compositions and methods useful for detecting, preventing,
 CC prognosing and/ or treating cancers that express 98P4B6. Furthermore, the
 CC polynucleotide can be used for gene therapy purposes, for monitoring,
 CC genetic abnormalities and for generating transgenic or 'knockout' CC
 CC animals, which can be useful for the development and screening of CC
 CC therapeutically useful reagents. The encoded proteins may also be used in CC
 CC generating antibodies and vaccines, as well as in identifying ligands and CC
 CC cellular constituents that bind to 98P4B6 gene products. Accordingly, CC
 CC these compositions exhibit cytostatic activities. This polynucleotide CC
 CC sequence is a partial human 98P4B6 cDNA variant of the invention. CC
 XX
 SQ Sequence 6857 BP; 2018 A; 1318 C; 1210 G; 2311 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
 Best Local Similarity 55.9%; Pred. No. 5.1;
 Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
 QY 41 TTGAAATATCTTCATTTGAAATAAACAGCTAACTCCCAATTTGCTAACTAGGAG 100
 DB 2234 TTGAATATTTTAAAGCTTAAATTTCTGCTGTTTGCATTTGTTTAACTATATC 2175
 QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATTAAGAGTAACAGAGTGTG 159
 DB 2174 TATATTAATTAAGTAAATTTTGGCTGATTAATGAATATAAGTAAACACAGGGTTA 2115
 QY 160 CAACATGCTGTTAAATACCGACAACTTCAATCACTATA 200
 DB 2114 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2074

RESULT 29
 ADW53585/c
 ID ADW53585 standard; cDNA; 6857 BP.
 XX
 AC ADW53585;
 XX
 DT 24-MAR-2005 (first entry)
 XX
 XX Human 98P4B6 cDNA variant 38 SeqID 76.
 DE ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
 KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
 KW cytostatic; gene; chromosome 7q21.
 XX
 OS Homo sapiens.
 XX
 XX WO2003087306-A2.
 XX
 XX 23-OCT-2003.
 XX
 XX 04-APR-2003; 2003WO-US010462.
 XX
 XX 05-APR-2002; 2002US-0370387P.
 XX
 XX 06-JUN-2002; 2002US-00165044.
 XX
 XX 20-DEC-2002; 2002US-0435480P.
 XX
 XX (AGEN-) AGENYSYS INC.
 XX
 XX Challita-Eid PM, Raitano AB, Paris M, Ge W, Jakobovits A;
 XX
 XX WPI; 2003-903158/82.
 XX
 XX P-PSDB; ADW53586.

A composition comprising 98P4B6 proteins and nucleic acid molecules for
 PT detecting, preventing, prognosing and/or treating cancers that express
 PT 98P4B6 (e.g. pancreatic, prostatic or lung cancer).
 XX
 XX Example 2; SEQ ID NO 76; 616pp; English.
 XX
 XX This invention relates to a novel human gene 98P4B6 (also known as STEAP-
 CC 2, six transmembrane epithelial antigen of prostate-2) located on

CC chromosome 7q21 and the encoded protein and variants derived thereof.
 CC Specifically, it refers to the expression pattern of this gene in adult
 CC normal tissues and its aberrant over-expression in various cancers
 CC including breast, colon, lung, kidney and prostate. The present invention
 CC describes compositions and methods useful for detecting, preventing,
 CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the
 CC polynucleotide can be used for gene therapy purposes, for monitoring
 CC genetic abnormalities and for generating transgenic or 'knockout'
 CC animals, which can be useful for the development and screening of
 CC therapeutically useful reagents. The encoded proteins may also be used in
 CC generating antibodies and vaccines, as well as in identifying ligands and
 CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
 CC these compositions exhibit cytostatic activities. This polynucleotide
 CC sequence is a human 98P4B6 cDNA variant of the invention.

XX Sequence 6857 BP; 2019 A; 1317 C; 1210 G; 2311 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
 Best Local Similarity 55.9%; Pred. No. 5.1;
 Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
 QY 41 TTGAATATATTTCTTCATTTGAAATTAACAGCTAACTCCCAAATTTGTACCTAACTAGGGGAG 100
 DB 2234 TTGAATATATTTTAAAGCTTAAATTTCTGCTGTTTGCATAATTTGTTTAACTATATC 2175
 QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTTCAATAAGAGTAACAGAGTGTG 159
 DB 2174 TATATTTAACTGTAAGTATTTTTCCTGCTGATTATGAATATATAAGTAACACAGGGTTA 2115
 QY 160 CAACATGCTGTTAAATAACCGACAACTTCAATCACTATA 200
 DB 2114 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2074

RESULT 30
 ADW53559/C
 ID ADW53559 standard; cDNA; 6857 BP.

XX ADW53559;
 XX 24-MAR-2005 (first entry)
 XX Human 98P4B6 cDNA variant 25 SeqID 50.
 XX ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
 XX prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
 XX cytostatic; gene; chromosome 7q21.

XX Homo sapiens.
 XX WO2003087306-A2.
 XX 23-OCT-2003.
 XX 04-APR-2003; 2003WO-US010462.
 XX 05-APR-2002; 2002US-0370387P.
 XX 06-JUN-2002; 2002US-00165044.
 XX 20-DEC-2002; 2002US-0435480P.
 XX (AGEN-) AGENYSYS INC.

XX Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
 XX WPI; 2003-903158/82.
 XX P-PSDB; ADW53560, ADW53596.
 XX A composition comprising 98P4B6 proteins and nucleic acid molecules for
 XX detecting, preventing, prognosing and/or treating cancers that express
 XX 98P4B6 (e.g. pancreatic, prostatic or lung cancer).
 XX Example 2; SEQ ID NO 50; 616pp; English.

CC This invention relates to a novel human gene 98P4B6 (also known as STEAP-
 CC 2, six transmembrane epithelial antigen of prostate-2) located on
 CC chromosome 7q21 and the encoded protein and variants derived thereof.
 CC Specifically, it refers to the expression pattern of this gene in adult
 CC normal tissues and its aberrant over-expression in various cancers
 CC including breast, colon, lung, kidney and prostate. The present invention
 CC describes compositions and methods useful for detecting, preventing,
 CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the
 CC polynucleotide can be used for gene therapy purposes, for monitoring
 CC genetic abnormalities and for generating transgenic or 'knockout'
 CC animals, which can be useful for the development and screening of
 CC therapeutically useful reagents. The encoded proteins may also be used in
 CC generating antibodies and vaccines, as well as in identifying ligands and
 CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
 CC these compositions exhibit cytostatic activities. This polynucleotide
 CC sequence is a human 98P4B6 cDNA variant of the invention.

XX Sequence 6857 BP; 2019 A; 1318 C; 1209 G; 2311 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
 Best Local Similarity 55.9%; Pred. No. 5.1;
 Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
 QY 41 TTGAATATATTTCTTCATTTGAAATTAACAGCTAACTCCCAAATTTGTACCTAACTAGGGGAG 100
 DB 2234 TTGAATATATTTTAAAGCTTAAATTTCTGCTGTTTGCATAATTTGTTTAACTATATC 2175
 QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTTCAATAAGAGTAACAGAGTGTG 159
 DB 2174 TATATTTAACTGTAAGTATTTTTCCTGCTGATTATGAATATATAAGTAACACAGGGTTA 2115
 QY 160 CAACATGCTGTTAAATAACCGACAACTTCAATCACTATA 200
 DB 2114 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2074

RESULT 31
 ADW53563/C
 ID ADW53563 standard; cDNA; 6857 BP.

XX ADW53563;
 XX 24-MAR-2005 (first entry)
 XX Human 98P4B6 cDNA variant 27 SeqID 54.
 XX ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
 XX prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
 XX cytostatic; gene; chromosome 7q21.

XX Homo sapiens.
 XX WO2003087306-A2.
 XX 23-OCT-2003.
 XX 04-APR-2003; 2003WO-US010462.
 XX 05-APR-2002; 2002US-0370387P.
 XX 06-JUN-2002; 2002US-00165044.
 XX 20-DEC-2002; 2002US-0435480P.
 XX (AGEN-) AGENYSYS INC.

XX Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
 XX WPI; 2003-903158/82.
 XX P-PSDB; ADW53564.
 XX A composition comprising 98P4B6 proteins and nucleic acid molecules for
 XX detecting, preventing, prognosing and/or treating cancers that express
 XX 98P4B6 (e.g. pancreatic, prostatic or lung cancer).

Example 2; SEQ ID NO 54; 616pp; English.

This invention relates to a novel human gene 98P4B6 (also known as STEAP-2, six transmembrane epithelial antigen of prostate-2) located on chromosome 7q21 and the encoded protein and variants derived thereof. Specifically it refers to the expression pattern of this gene in adult normal tissues and its aberrant over-expression in various cancers including breast, colon, lung, kidney and prostate. The present invention describes compositions and methods useful for detecting, preventing, prognosing and/or treating cancers that express 98P4B6. Furthermore, the polynucleotide can be used for gene therapy purposes, for monitoring genetic abnormalities and for generating transgenic or 'knockout' animals, which can be useful for the development and screening of therapeutically useful reagents. The encoded proteins may also be used in generating antibodies and vaccines, as well as in identifying ligands and cellular constituents that bind to 98P4B6 gene products. Accordingly, these compositions exhibit cytostatic activities. This polynucleotide sequence is a human 98P4B6 cDNA variant of the invention.

Sequence 6857 BP; 2018 A; 1317 C; 1210 G; 2312 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

41 TTGAAATATCTTCAATTTGAAATAAACAAGCTAACTCCCAATTTGACCTAAGTAGGGAG 100
|||||
2234 TTGAATATTTTAAAGCTTAAATTCGCTGGTTTGCATAATGTTTAAACATTATATC 2175
|||||
101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAGTAACAGAGTGTG 159
|||||
2174 TATATTATTAAGCTGAAGTATTTTTCCTGCTGATTATGAATAATAAGTAACAACAGGGTTA 2115
|||||
160 CAACATGCTGTTAAATAACCCGACAACTTCAATCACTATA 200
|||||
2114 CAGTTGTGCAAAACAAACAAACATACCTTAATTAACAATA 2074
|||||

RESULT 32
ADW53525/c
ID ADW53525 standard; cDNA; 6857 BP.
XX ADW53525;
AC ADW53525;
XX 24-MAR-2005 (first entry)
XX Human 98P4B6 cDNA variant 8 SeqID 16.
XX ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
KW cytosolic; gene; chromosome 7q21.
XX Homo sapiens.
XX WO2003087306-A2.
XX 23-OCT-2003.
XX 04-APR-2003; 2003WO-US010462.
XX 05-APR-2002; 2002US-0370387P.
PR 06-JUN-2002; 2002US-00165044.
PR 20-DEC-2002; 2002US-0435480P.
XX (AGEN-) AGENYSYS INC.
XX Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
XX WPI; 2003-903158/82.
XX P-PSDB; ADW53526, ADW53592.
XX A composition comprising 98P4B6 proteins and nucleic acid molecules for
PT detecting, preventing, prognosing and/or treating cancers that express

Example 2; SEQ ID NO 16; 616pp; English.

This invention relates to a novel human gene 98P4B6 (also known as STEAP-2, six transmembrane epithelial antigen of prostate-2) located on chromosome 7q21 and the encoded protein and variants derived thereof. Specifically it refers to the expression pattern of this gene in adult normal tissues and its aberrant over-expression in various cancers including breast, colon, lung, kidney and prostate. The present invention describes compositions and methods useful for detecting, preventing, prognosing and/or treating cancers that express 98P4B6. Furthermore, the polynucleotide can be used for gene therapy purposes, for monitoring genetic abnormalities and for generating transgenic or 'knockout' animals, which can be useful for the development and screening of therapeutically useful reagents. The encoded proteins may also be used in generating antibodies and vaccines, as well as in identifying ligands and cellular constituents that bind to 98P4B6 gene products. Accordingly, these compositions exhibit cytostatic activities. This polynucleotide sequence is a human 98P4B6 cDNA variant of the invention.

Sequence 6857 BP; 2018 A; 1318 C; 1210 G; 2311 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
Best Local Similarity 55.9%; Pred. No. 5.1;
Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

41 TTGAAATATCTTCAATTTGAAATAAACAAGCTAACTCCCAATTTGACCTAAGTAGGGAG 100
|||||
2234 TTGAATATTTTAAAGCTTAAATTCGCTGGTTTGCATAATGTTTAAACATTATATC 2175
|||||
101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAGTAACAGAGTGTG 159
|||||
2174 TATATTATTAAGCTGAAGTATTTTTCCTGCTGATTATGAATAATAAGTAACAACAGGGTTA 2115
|||||
160 CAACATGCTGTTAAATAACCCGACAACTTCAATCACTATA 200
|||||
2114 CAGTTGTGCAAAACAAACAAACATACCTTAATTAACAATA 2074
|||||

RESULT 33
ADW53565/c
ID ADW53565 standard; cDNA; 6857 BP.
XX ADW53565;
AC ADW53565;
XX 24-MAR-2005 (first entry)
XX Human 98P4B6 cDNA variant 28 SeqID 56.
XX ss; human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
KW cytosolic; gene; chromosome 7q21.
XX Homo sapiens.
XX WO2003087306-A2.
XX 23-OCT-2003.
XX 04-APR-2003; 2003WO-US010462.
XX 05-APR-2002; 2002US-0370387P.
PR 06-JUN-2002; 2002US-00165044.
PR 20-DEC-2002; 2002US-0435480P.
XX (AGEN-) AGENYSYS INC.
XX Challita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;
XX WPI; 2003-903158/82.
XX P-PSDB; ADW53566.
XX A composition comprising 98P4B6 proteins and nucleic acid molecules for
PT detecting, preventing, prognosing and/or treating cancers that express

PT A composition comprising 98P4B6 proteins and nucleic acid molecules for
 PT detecting, preventing, prognosing and/or treating cancers that express
 XX 98P4B6 (e.g. pancreatic, prostatic or lung cancer).
 XX Example 2; SEQ ID NO 56; 616pp; English.
 XX This invention relates to a novel human gene 98P4B6 (also known as STEAP-
 CC 2, six transmembrane epithelial antigen of prostate-2) located on
 CC chromosome 7q21 and the encoded protein and variants derived thereof.
 CC Specifically, it refers to the expression pattern of this gene in adult
 CC normal tissues and its aberrant over-expression in various cancers
 CC including breast, colon, lung, kidney and prostate. The present invention
 CC describes compositions and methods useful for detecting, preventing,
 CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the
 CC polynucleotide can be used for gene therapy purposes, for monitoring
 CC genetic abnormalities and for generating transgenic or 'knockout'
 CC animals, which can be useful for the development and screening of
 CC therapeutically useful reagents. The encoded proteins may also be used in
 CC generating antibodies and vaccines, as well as in identifying ligands and
 CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
 CC these compositions exhibit cytostatic activities. This polynucleotide
 CC sequence is a human 98P4B6 cDNA variant of the invention.
 XX
 SQ Sequence 6857 BP; 2018 A; 1319 C; 1210 G; 2310 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
 Best Local Similarity 55.9%; Pred. No. 5.1;
 Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
 QY 41 TTGAATATTTCTTCATTTGAAATGAAACAGCTAACTCCCAATTTGCTAACTAAGGGAG 100
 DB 2234 TTGAATATTTTAAAGCTTAAATCTGCTGCTTTCGAAATTTGTTTAACTATATC 2175
 QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTCAAATAAGAGTAACACAGAGTGTG 159
 DB 2174 TATATTATTAAGTAACTGTAAGTATTTTGGCTGATTGAAATATAAGTAAACACAGGTTA 2115
 QY 160 CAACATGCTGTTAAATTAACCCGACAAACTTCAATCACTATA 200
 DB 2114 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2074

RESULT 34
 ADW53567/c
 ID: ADW53567 standard; cDNA; 6857 BP.

XX: ADW53567;

XX: 24-MAR-2005 (first entry)

XX: Human 98P4B6 cDNA variant 29 SeqID 58.

DE ss: human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney;
 KW prostate; gene therapy; genetic abnormality; transgenic; knockout animal;
 KW Cytostatic; gene; chromosome 7q21.

XX: Homo sapiens.

XX: WO2003087306-A2.

XX: 23-OCT-2003.

XX: 04-APR-2003; 2003WO-US010462.

XX: 05-APR-2002; 2002US-0370387P.

XX: 06-JUN-2002; 2002US-00165044.

XX: 20-DEC-2002; 2002US-0435480P.

XX: (AGEN-) AGENYSYS INC.

XX: Challita-Eid PM, Raitano AB, Paris M, Ge W, Jakobovits A;

XX: WPI; 2003-903158/82.

DR P-PSDB; ADW53568.
 XX A composition comprising 98P4B6 proteins and nucleic acid molecules for
 PT detecting, preventing, prognosing and/or treating cancers that express
 XX 98P4B6 (e.g. pancreatic, prostatic or lung cancer).
 XX Example 2; SEQ ID NO 58; 616pp; English.
 XX This invention relates to a novel human gene 98P4B6 (also known as STEAP-
 CC 2, six transmembrane epithelial antigen of prostate-2) located on
 CC chromosome 7q21 and the encoded protein and variants derived thereof.
 CC Specifically, it refers to the expression pattern of this gene in adult
 CC normal tissues and its aberrant over-expression in various cancers
 CC including breast, colon, lung, kidney and prostate. The present invention
 CC describes compositions and methods useful for detecting, preventing,
 CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the
 CC polynucleotide can be used for gene therapy purposes, for monitoring
 CC genetic abnormalities and for generating transgenic or 'knockout'
 CC animals, which can be useful for the development and screening of
 CC therapeutically useful reagents. The encoded proteins may also be used in
 CC generating antibodies and vaccines, as well as in identifying ligands and
 CC cellular constituents that bind to 98P4B6 gene products. Accordingly,
 CC these compositions exhibit cytostatic activities. This polynucleotide
 CC sequence is a human 98P4B6 cDNA variant of the invention.
 XX
 SQ Sequence 6857 BP; 2019 A; 1318 C; 1210 G; 2310 T; 0 U; 0 Other;

Query Match 12.3%; Score 37; DB 11; Length 6857;
 Best Local Similarity 55.9%; Pred. No. 5.1;
 Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
 QY 41 TTGAATATTTCTTCATTTGAAATGAAACAGCTAACTCCCAATTTGCTAACTAAGGGAG 100
 DB 2234 TTGAATATTTTAAAGCTTAAATCTGCTGCTTTCGAAATTTGTTTAACTATATC 2175
 QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAAATTTCAAATAAGAGTAACACAGAGTGTG 159
 DB 2174 TATATTATTAAGTAACTGTAAGTATTTTGGCTGATTGAAATATAAGTAAACACAGGTTA 2115
 QY 160 CAACATGCTGTTAAATTAACCCGACAAACTTCAATCACTATA 200
 DB 2114 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2074

RESULT 35
 ADW66608/c

ID: ADW66608 standard; DNA; 6857 BP.

XX: ADW66608;

XX: 02-DEC-2004 (first entry)

XX: Human prostatic carcinoma derived DNA SEQ ID 120 #3.

XX human; cytostatic; diagnosis; prostatic cancer;
 KW differential expression analysis; ds.

XX: Homo sapiens.

XX: WO2004076614-A2.

XX: 10-SEP-2004.

XX: 22-FEB-2004; 2004WO-DE000433.

XX: 27-FEB-2003; 2003DE-0100985.

XX: 14-MAY-2003; 2003DE-0102134.

XX: (HINZ/) HINZMANN B.

XX: (DAHL/) DAHL E.

XX: (ROSE/) ROSENTHAL A.

XX: (HERM/) HERMANN K.

XX: (PILA/) PILARSKY C.

XX
PI Hinzmänn B, Dahl E, Rosenthal A, Hermann K, Pillarsky C, Specht T;
PI Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;
PI Xinzhang L, Staub E;
XX
XX MPI; 2004-653386/63.
XX
XX New nucleic acids, and encoded proteins, from prostatic cancer tissue,
XX useful for diagnosis, treatment and in screening for specific binding
XX agents.
XX
XX Claim 1; Page 984-986; 1607pp; German.
XX
XX This invention describes novel cytostatic polynucleotide and polypeptide
XX sequences which can be used in a method for diagnosing prostatic cancer
XX or the risk of developing prostatic cancer. Diagnosis is based on
XX determining over transcription or over expression of the sequences in
XX prostatic tissue. Screening for inhibitors of the sequences or detection
XX substances involves a binding assay, any compounds that bind are
XX selected, optionally after deconvolution of mixtures. Detection of a
XX predetermined minimum level of the reporter indicates the presence of
XX tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
XX short-interfering RNA or ribozymes; an organic molecule of molecular
XX weight below 5000, preferably 300, that binds to the polypeptide; an
XX aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
XX polypeptide, preferably humanised or human; an anti-idiotypic, non-human
XX (monoclonal) antibody directed against Ab or any of the above derivatised
XX with a reporter group, cell toxin, immunostimulatory molecules and/or
XX radioisotope. The polynucleotides are identified in human prostatic
XX cancer by differential expression analysis, using DNA microarrays,
XX between normal and tumorous tissues, with (over)expression being detected
XX by quantitative PCR. Analysis of prostatic cancer samples showed that
XX CD24 was upregulated in many of them. Sections of tissue, isolated from
XX prostatic cancer patients, or subjects at risk, were incubated
XX sequentially with anti-human CD4 murine monoclonal antibodies;
XX biotinylated second antibody; streptavidin-conjugated horseradish
XX peroxidase and then diaminobenzidine as colour former (brown). The
XX samples were counterstained with hemalum (blue). Malignant cells stained
XX strongly but non-malignant cells only weakly. In 15 of 63 samples of
XX adenocarcinoma, membrane and cytoplasmic staining was very strong, and
XX lymph node metastases were also stained. ADR65805-ADR66954 represent the
XX polynucleotide and polypeptide sequences used in the method of the
XX invention.
XX
XX Sequence 6857 BP; 2021 A; 1317 C; 1208 G; 2311 T; 0 U; 0 Other;
XX
XX Query Match 12.3%; Score 37; DB 13; Length 6857;
XX Best Local Similarity 55.9%; Pred. No. 5.1;
XX Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
XX
XX 41 TTGAATATTCTTCATTTGAAATAAACAGCTAACTCCCAATTTGCTACTACTAGGGAG 100
XX 2234 TTGAATATTGTTAAAGCTTAAATTCGTGGTTTGCATAATGTTTAACTATATATC 2175
XX
XX 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAGTAACAGAGTGTG 159
XX 2174 TATATTATTAATCTGTAAGTATTTTTCCTGATTGAAATATAAGTAACAACAGGGTTA 2115
XX
XX 160 CAACATGCTGTTTAATTAACCCGACAACTTCAATCACTATA 200
XX 2114 CAGTTGTGCAAAACAAACAAACATACCTTAATTAACAATA 2074
XX
XX
XX RESULT 36
XX ADR66266/C
XX ID ADR66266 standard; DNA; 6857 BP.
XX
XX ADR66266;
XX
XX 02-DEC-2004 (first entry)
XX
XX Human prostatic carcinoma derived DNA SEQ ID 120 #2.
XX

KW human; cytostatic; diagnosis; prostatic cancer;
KW differential expression analysis; ds.
XX
XX OS Homo sapiens.
XX
XX WO2004076614-A2.
XX
XX 10-SEP-2004.
XX
XX 22-FEB-2004; 2004WO-DE000433.
XX
XX 27-FEB-2003; 2003DE-01009985.
XX 14-MAY-2003; 2003DE-01022134.
XX
XX (HINZ/) HINZMANN B.
XX (DAHL/) DAHL E.
XX (ROSE/) ROSENTHAL A.
XX (HERM/) HERMANN K.
XX (PILA/) PILARSKY C.
XX
XX Hinzmänn B, Dahl E, Rosenthal A, Hermann K, Pillarsky C, Specht T;
XX Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;
XX Xinzhang L, Staub E;
XX
XX MPI; 2004-653386/63.
XX
XX New nucleic acids, and encoded proteins, from prostatic cancer tissue,
XX useful for diagnosis, treatment and in screening for specific binding
XX agents.
XX
XX Claim 1; Page 338-339; 1607pp; German.
XX
XX This invention describes novel cytostatic polynucleotide and polypeptide
XX sequences which can be used in a method for diagnosing prostatic cancer
XX or the risk of developing prostatic cancer. Diagnosis is based on
XX determining over transcription or over expression of the sequences in
XX prostatic tissue. Screening for inhibitors of the sequences or detection
XX substances involves a binding assay, any compounds that bind are
XX selected, optionally after deconvolution of mixtures. Detection of a
XX predetermined minimum level of the reporter indicates the presence of
XX tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
XX short-interfering RNA or ribozymes; an organic molecule of molecular
XX weight below 5000, preferably 300, that binds to the polypeptide; an
XX aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
XX polypeptide, preferably humanised or human; an anti-idiotypic, non-human
XX (monoclonal) antibody directed against Ab or any of the above derivatised
XX with a reporter group, cell toxin, immunostimulatory molecules and/or
XX radioisotope. The polynucleotides are identified in human prostatic
XX cancer by differential expression analysis, using DNA microarrays,
XX between normal and tumorous tissues, with (over)expression being detected
XX by quantitative PCR. Analysis of prostatic cancer samples showed that
XX CD24 was upregulated in many of them. Sections of tissue, isolated from
XX prostatic cancer patients, or subjects at risk, were incubated
XX sequentially with anti-human CD4 murine monoclonal antibodies;
XX biotinylated second antibody; streptavidin-conjugated horseradish
XX peroxidase and then diaminobenzidine as colour former (brown). The
XX samples were counterstained with hemalum (blue). Malignant cells stained
XX strongly but non-malignant cells only weakly. In 15 of 63 samples of
XX adenocarcinoma, membrane and cytoplasmic staining was very strong, and
XX lymph node metastases were also stained. ADR65805-ADR66954 represent the
XX polynucleotide and polypeptide sequences used in the method of the
XX invention.
XX
XX Sequence 6857 BP; 2021 A; 1317 C; 1208 G; 2311 T; 0 U; 0 Other;
XX
XX Query Match 12.3%; Score 37; DB 13; Length 6857;
XX Best Local Similarity 55.9%; Pred. No. 5.1;
XX Matches 90; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
XX
XX 41 TTGAATATTCTTCATTTGAAATAAACAGCTAACTCCCAATTTGCTACTACTAGGGAG 100
XX 2234 TTGAATATTGTTAAAGCTTAAATTCGTGGTTTGCATAATGTTTAACTATATATC 2175
XX
XX 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAGTAACAGAGTGTG 159
XX 2174 TATATTATTAATCTGTAAGTATTTTTCCTGATTGAAATATAAGTAACAACAGGGTTA 2115
XX
XX 160 CAACATGCTGTTTAATTAACCCGACAACTTCAATCACTATA 200
XX 2114 CAGTTGTGCAAAACAAACAAACATACCTTAATTAACAATA 2074
XX
XX
XX RESULT 36
XX ADR66266/C
XX ID ADR66266 standard; DNA; 6857 BP.
XX
XX ADR66266;
XX
XX 02-DEC-2004 (first entry)
XX
XX Human prostatic carcinoma derived DNA SEQ ID 120 #2.
XX

Claim 8; SEQ ID NO 17892; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides, and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

Sequence 2272 BP; 734 A; 333 C; 347 G; 858 T; 0 U; 0 Other;

Query Match 12.3%; Score 36.8; DB 4; Length 2272;
Best Local Similarity 50.4%; Pred. No. 4.3;
Matches 121; Conservative 0; Mismatches 112; Indels 7; Gaps 1;

44 AATATCTCTCAATTTGAATTAACACAGCTAACTCCAAATTTACTACTAGGGAGTTT 103
1281 AGATATTTTAAATTAACACACATTTTCTTTAAAGGTCATATCTATTAGGCCTT 1340
104 ATCATGAGAAATCTTTAAATGCTTAAATTTCAATTAAGAGTAACAGAGTGTTCAC 163
1341 TCCCTGCTGTAATCTTAAATGATTTTGTGAATTTTACTAA-----GTTTCATC 1393
164 ATGCTGTTAAATACCCGACAACTTCAATCACTATAGCTCTAGTAGAGTGCATTTCTGCA 223
1394 AAGATCTTAAAGAGAGAGAGAAATAAAGATATTATTATATGTCAGTCCCACTTTG 1453
224 AGGATCCAGAGTAACAGATGTTTGGAAATGCAATGTTGAACCGACCATTAATATC 283
1454 AAGGTAAACAGTAGTAGTAACTTTTGGGAGAAATCTTAAAAACTACTTTAAAAAAGAC 1513

RESULT 39

AAK84697
ID AAK84697 standard; DNA; 2533 BP.

AAK84697;

07-NOV-2001 (first entry)

Human immune/haematopoietic antigen genomic sequence SEQ ID NO:39509.

Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
cytostatic; gene therapy; vaccine; metastasis; ds.

Homo sapiens.

WO200157182-A2.

09-AUG-2001.

17-JAN-2001; 2001WO-US001354.

31-JAN-2000; 2000US-0179065P.

04-FEB-2000; 2000US-0180628P.

24-FEB-2000; 2000US-0184664P.

02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231988P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233083P.
PR 14-SEP-2000; 2000US-0233084P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235836P.
PR 27-SEP-2000; 2000US-0235837P.
PR 29-SEP-2000; 2000US-0236372P.
PR 29-SEP-2000; 2000US-0236377P.
PR 29-SEP-2000; 2000US-0236378P.
PR 29-SEP-2000; 2000US-0236386P.
PR 29-SEP-2000; 2000US-0236389P.
PR 29-SEP-2000; 2000US-0236390P.
PR 02-OCT-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.

PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225211P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226689P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0227009P.
PR 01-SEP-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.

Disclosure; SEQ ID NO 39510; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the

XX PI Lesche R, Fassbender A, Juenemann K, Foekens J, Martens JMW;
 XX DR WPI; 2006-125479/13.
 XX PT Detecting or detecting and distinguishing between or among breast cell
 XX PT proliferative disorders in a subject by determining the expression of a
 XX PT panel of genes or sequences.
 XX PS Claim 14; SEQ ID NO 913; 321pp; English.
 XX CC The invention relates to a method of detecting or detecting and
 XX CC distinguishing between or among breast cell proliferative disorders in a
 XX CC subject comprising determining the expression of at least one gene or
 XX CC sequence given in the specification, and determining from the expression
 XX CC the presence, absence or subclass of a breast cell proliferative
 XX CC disorder. The invention also relates to a treated nucleic acid used to
 XX CC convert at least one unmethylated cytosine base of the genomic DNA
 XX CC sequence to uracil or another base that is detectably dissimilar to
 XX CC cytosine in terms of hybridization, and an oligomer that is complementary
 XX CC to, or hybridizes under moderately stringent conditions to a treated
 XX CC genomic DNA sequence or a set of oligomers comprising at least two
 XX CC oligonucleotides. The method is useful for detecting or detecting and
 XX CC distinguishing between or among breast cell proliferative disorders in a
 XX CC subject. This sequence represents human genomic DNA used in the scope of
 XX CC the invention. Note: The sequence data for this patent did not form part
 XX CC of the printed specification, but was obtained in electronic format
 XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 2374 BP; 536 A; 0 C; 675 G; 1163 T; 0 U; 0 Other;
 Query Match 12.1%; Score 36.4; DB 15; Length 2374;
 Best Local Similarity 47.1%; Pred. No. 5.7;
 Matches 112; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
 QY 37 AGACTGGAATATCTTCATTTGAAATAAACAGCTAACTCCCAATTTGTACCTAACTAGG 96
 Db 275 AAAATTTAAATATATATTTTAAATTAACAAAACACTTAAATTTAAATTAATCTATAT 216
 QY 97 GGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAATTAAGAAAGTAACCCAGAGTG 156
 Db 215 CACTTTTTTAAATACATATATTTAACTTAAATTAATTAATTAATTAATTAATTAATTT 156
 QY 157 TTGCAACATGCTGTTAAATTAACCCGACAACTTCAATCACTATAGCTGTAGTAGTGCA 216
 Db 155 AAATAAAATCATATATACTTAAACAACTCTAATATATAATAATACTATAAAAAACGA 96
 QY 217 TTCTGCAAGGATCCGAGTAACAGTATTTTGGAAATGCAATGTTGAAACCGACATA 274
 Db 95 ACACACAAAACATAATAATAACAAAATTAATAATAAAACTATATTTATTACTAACTCTA 38
 RESULT 45
 AEF55696/c
 ID AEF55696 standard; DNA; 2374 BP.
 XX AC AEF55696;
 XX AC AEF55696;
 XX DT 23-MAR-2006 (first entry)
 XX DE Human genomic DNA #303.
 XX KW Diagnosis; breast disease; hyperproliferation; gene; ds.
 XX OS Homo sapiens.
 XX EN WO2006008128-A2.
 XX FX 26-JAN-2006.
 XX PD 18-JUL-2005; 2005WO-EP007830.
 XX PF 18-JUL-2004; 2004EP-00016926.
 XX PR

XX PA (BPIG-) EPIGENOMICS AG.
 XX PI Lesche R, Fassbender A, Juenemann K, Foekens J, Martens JMW;
 XX DR WPI; 2006-125479/13.
 XX PT Detecting or detecting and distinguishing between or among breast cell
 XX PT proliferative disorders in a subject by determining the expression of a
 XX PT panel of genes or sequences.
 XX PS Claim 14; SEQ ID NO 677; 321pp; English.
 XX CC The invention relates to a method of detecting or detecting and
 XX CC distinguishing between or among breast cell proliferative disorders in a
 XX CC subject comprising determining the expression of at least one gene or
 XX CC sequence given in the specification, and determining from the expression
 XX CC the presence, absence or subclass of a breast cell proliferative
 XX CC disorder. The invention also relates to a treated nucleic acid used to
 XX CC convert at least one unmethylated cytosine base of the genomic DNA
 XX CC sequence to uracil or another base that is detectably dissimilar to
 XX CC cytosine in terms of hybridization, and an oligomer that is complementary
 XX CC to, or hybridizes under moderately stringent conditions to a treated
 XX CC genomic DNA sequence or a set of oligomers comprising at least two
 XX CC oligonucleotides. The method is useful for detecting or detecting and
 XX CC distinguishing between or among breast cell proliferative disorders in a
 XX CC subject. This sequence represents human genomic DNA used in the scope of
 XX CC the invention. Note: The sequence data for this patent did not form part
 XX CC of the printed specification, but was obtained in electronic format
 XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 2374 BP; 536 A; 146 C; 675 G; 1017 T; 0 U; 0 Other;
 Query Match 12.1%; Score 36.4; DB 15; Length 2374;
 Best Local Similarity 47.1%; Pred. No. 5.7;
 Matches 112; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
 QY 37 AGACTGGAATATCTTCATTTGAAATAAACAGCTAACTCCCAATTTGTACCTAACTAGG 96
 Db 275 AAAATTTAAATATATATTTTAAATTAACAAAACACTTAAATTTAAATTAATCTATAT 216
 QY 97 GGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAATTAAGAAAGTAACCCAGAGTG 156
 Db 215 CACTTTTTTAAATACATATATTTAACTTAAATTAATTAATTAATTAATTAATTT 156
 QY 157 TTGCAACATGCTGTTAAATTAACCCGACAACTTCAATCACTATAGCTGTAGTAGTGCA 216
 Db 155 AAATAAAATCATATATACTTAAACAACTCTAATATATAATAATACTATAAAAAACGA 96
 QY 217 TTCTGCAAGGATCCGAGTAACAGTATTTTGGAAATGCAATGTTGAAACCGACATA 274
 Db 95 ACACACAAAACATAATAATAACAAAATTAATAATAAAACTATATTTATTACTAACTCTA 38
 RESULT 46
 ACL35363
 ID ACL35363 standard; cDNA; 2000 BP.
 XX AC ACL35363;
 XX AC ACL35363;
 XX DT 02-JUN-2005 (first entry)
 XX DE Rice stress-regulated promoter SEQ ID NO:13926.
 XX KW ss; abiotic stress tolerance; transgenic plant; plant; cereal;
 XX KW agriculture.
 XX OS Oryza sativa.
 XX EN WO2003008540-A2.
 XX FX 30-JAN-2003.
 XX PD
 XX PR

21-JUN-2002; 2002WO-US019668.
 22-JUN-2001; 2001US-0300112P.
 24-AUG-2001; 2001US-0314662P.
 26-SEP-2001; 2001US-0325277P.
 21-NOV-2001; 2001US-0332132P.
 (SYGN) SYNGENTA PARTICIPATIONS AG.
 Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F, Moughamer T, Provart N, Ricke D, Zhu T, WPI; 2003-248011/24.
 New stress-responsive nucleic acid, useful for altering the responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold stress, salt stress or osmotic stress.
 Claim 48; SEQ ID NO 13926; 89pp; English.
 The invention relates to novel abiotic stress responsive polynucleotides and polypeptides. Also disclosed are vectors, expression cassettes, host cells, and plants containing such polynucleotides. Also disclosed are methods for using the polynucleotides and polypeptides to alter the responsiveness of a plant to abiotic stress. The invention is useful in agriculture. The nucleic acid is useful for determining whether a test plant has been exposed to an abiotic stress condition. It is also useful for selecting an agent that alters abiotic stress regulated polynucleotide expression in a plant cell, and to identify a homolog or ortholog to an abiotic stress responsive polynucleotide. The nucleic acid molecule and the polypeptide encoded by it are useful in altering the responsiveness of a plant to an abiotic stress, such as cold stress, salt stress, osmotic stress or any of their combinations. The present sequence is used in the exemplification of the invention
 Sequence 2000 BP; 540 A; 344 C; 416 G; 428 T; 0 U; 272 Other;
 Query Match 12.1%; Score 36.2; DB 11; Length 2000;
 Best Local Similarity 12.3%; Pred. No. 6.1;
 Matches 29; Conservative 109; Mismatches 97; Indels 0; Gaps 0;
 17 ATGCTAAGTACGATGACGAGCTGTAATATTTCTTCAATTTGAATAACAGCTAATC 76
 123 WTRMYNNKACTCYCTMYTMYRWAAARWKKWYKMWARGWSAKMKAYRYKWCYTY 182
 77 CCAAAATGTACTTAACATGCGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCA 136
 183 WAWMAWRWMTTKCTNSWYAWNTCYATWMAWATAWTTTTRMAWKKTRAWTTWRYAW 242
 137 AATAAGAAGTAACGAGAGTGTGCAACATGCTGTGTAATAAACCACCAAACTTCAATCAC 196
 243 MYTWWAKWKWYATKWTYAWRWMMWMMMSRTYRWMTWAWMYTWSWYTTWALMSWY 302
 197 TATAGCTGTAGTAGAGTCAATCTGCAAGGATCCGAGAGTACCAAGTATTTTGGGA 251
 303 WWCMTTWSRRRWYRKSRRTWTWTWTWMAWKKWYKMWMAWMAWKGW 357
 RESULT 47
 ABL70238/c
 ID ABL70238 standard; DNA; 9087 BP.
 AC ABL70238;
 23-APR-2002 (first entry)
 Signal transduction associated gene modified complementary DNA #60.
 Human; signal transduction associated gene; cytosine methylation state; CpG island; signal transduction associated disease; solid tumour; cancer; antitumour; cytosine; mutant; ds.
 Homo sapiens.

XX DE Chemically treated cell signalling DNA sequence complementary to#64.
 XX KW Cell signalling; cytosine methylation; cell signalling disease; cancer;
 XX KW tumour; cytostatic; ds.
 XX OS Unidentified.
 XX PN WO200202807-A2.
 XX PD 10-JAN-2002.
 XX XX 29-JUN-2001; 2001WO-EP007471.
 XX PR 30-JUN-2000; 2000DE-01032529.
 XX PR 01-SEP-2000; 2000DE-01043826.
 XX XX (EPIG-) EPIGENOMICS AG.
 XX PI Olek A, Piepenbrock C, Berlin K;
 XX XX WPI; 2002-154758/20.
 XX XX Nucleic acid, useful for diagnosis and therapy of diseases associated
 PT with cell signaling e.g. cancer, comprises chemically modified genomic
 PT sequences of genes associated with cell signaling.
 XX XX Claim 1; SEQ ID NO 128; 24pp + Sequence Listing; English.
 XX XX The invention relates to a nucleic acid comprising a sequence of at least
 CC 18 bases of a segment of chemically pretreated DNA of genes associated
 CC with cell signalling. The activity of the modified sequences of the
 CC invention may be described as cytostatic. The object of the invention is
 CC to provide the chemically modified DNA of genes associated with cell
 CC signalling, as well as oligonucleotides and/or PNA-oligonucleotides for
 CC detecting cytosine methylations, as well as a method which is
 CC particularly suitable for the diagnosis and/or therapy of genetic and
 CC epigenetic parameters of genes associated with cell signalling. The
 CC chemically modified DNA provided by the invention is useful for diagnosis
 CC and therapy of diseases such as solid tumours and cancer. The sequences
 CC given in records ABL70111-ABL70626 represent chemically pre-treated
 CC genomic DNA's of genes associated with cell signalling. Note: The
 CC sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 XX European Patent Office
 XX SQ Sequence 9087 BP; 2207 A; 253 C; 2592 G; 4034 T; 0 U; 1 Other;
 Query Match 12.1%; Score 36.2; DB 6; Length 9087;
 Best Local Similarity 48.3%; Pred. No. 9.1;
 Matches 101; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
 QY 79 AAATTGTACCTAACTAGGGAGTTTATCATCAAGAAATGTTTAAATGCTTAATTTTCAA 138
 DB 1413 AAAATTCCTTCTAAAAAACCTTACCTTAATTAATAATAATATCTATTACTACT 1354
 QY 139 TAAGAAGTAACAGAGTGTTCGACATGCTGTTAAATAACCCGACAACTTCAATCACTA 198
 DB 1353 CACAATAAACAATAACTACAAAACATATTCGCAAAAACCTCTAAACAAATTTCTCAATA 1294
 QY 199 TAGCTGTAGTAGAGTGCAATTCGCAAGATCCGAGATGACAGTATTTTGGAAATGCAA 258
 DB 1293 AACCCCTACATACTCTCTCTCCAAACAACCTTAATCGAAAATATCTATAAAACCTAA 1234
 QY 259 TGTGTGAACCGACCATACTAAATATCTGCT 287
 DB 1233 CCTTACCCCTCCCTTAAAAACAATCTTCT 1205
 RESULT 49
 AAS61181/c
 ID AAS61181 standard; DNA; 9087 BP.
 XX

AC AAS61181;
 XX 29-JAN-2002 (first entry)
 XX DE Human gene regulation-associated gene oligonucleotide #136.
 XX XX Human; Gene regulation-associated gene; severe combined immunodeficiency;
 KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
 KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
 KW renal disease; Preclampsia; cardiac allograft vascular disease;
 KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
 KW immunostimulant; cardiant; antiinflammatory; coagulant; antiasthmatic;
 KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.
 XX OS Homo sapiens.
 XX XX WO200177375-A2.
 XX PN 18-OCT-2001.
 XX PD 06-APR-2001; 2001WO-EP003968.
 XX PF 06-APR-2000; 2000DE-01019058.
 XX PR 07-APR-2000; 2000DE-01019173.
 XX PR 30-JUN-2000; 2000DE-01032529.
 XX PR 01-SEP-2000; 2000DE-01043826.
 XX XX (EPIG-) EPIGENOMICS AG.
 XX PA Olek A, Piepenbrock C, Berlin K;
 XX PI WPI; 2002-017470/02.
 XX DR New nucleic acid sequences from chemically modified genes associated with
 PT gene regulation, useful for analyzing cytosine methylations for diagnosis
 PT and therapy of diseases e.g. severe combined immunodeficiency disease.
 XX XX Claim 1; SEQ ID NO 140; 26pp; English.
 XX PS The invention relates to 224 nucleic acid sequences comprising at least
 XX 18 bases of a chemically pretreated gene associated with gene regulation
 CC selected from 43 known genes (or complementary sequences). The chemical
 CC pretreatment converts cytosine bases unmethylated at the 5-position to
 CC uracil or another base with hybridisation behaviour dissimilar to
 CC cytosine, to enable analysis of cytosine methylations. The DNA sequences
 CC oligomers (or sets/arrays) and method are useful in the diagnosis of
 CC diseases (or predisposition to diseases) associated with gene regulation
 CC and in therapy of such diseases, by enabling analysis of the cytosine
 CC methylation patterns of such genes, kits are provided. They are
 CC especially useful in diagnosis and therapy of e.g. severe combined
 CC immunodeficiency disease, cardiac disorders, haemophilia, solid tumours
 CC and cancer, Werner syndrome, asthma, HDR syndrome, Saethre-Chotzen
 CC syndrome, renal disease, preclampsia, graft versus-host disease. The
 CC present sequence is a sequence included in the sequence data for this
 CC specification and is associated with the human gene regulation-associated
 CC genes. Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 9087 BP; 2207 A; 253 C; 2592 G; 4034 T; 0 U; 1 Other;
 Query Match 12.1%; Score 36.2; DB 6; Length 9087;
 Best Local Similarity 48.3%; Pred. No. 9.1;
 Matches 101; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
 QY 79 AAATTGTACCTAACTAGGGAGTTTATCATCAAGAAATGTTTAAATGCTTAATTTTCAA 138
 DB 1413 AAAATTCCTTCTAAAAAACCTTACCTTAATTAATAATAATATCTATTACTACT 1354
 QY 139 TAAGAAGTAACAGAGTGTTCGACATGCTGTTAAATAACCCGACAACTTCAATCACTA 198
 DB 1353 CACAATAAACAATAACTACAAAACATATTCGCAAAAACCTCTAAACAAATTTCTCAATA 1294

Qy 199 TAGCTGTAGTAGTGCATTCTGCAAGGATCCGAGTAGTAACAGTATTTCGGAATGCAA 258
Db 1293 AACCCCTACAATCTCTCTTCCAAACAATTAATCGAAAAATCATCTAAAAACCTAA 1234
Qy 259 TGTTGAACCGACCATACTAATTAATCTGCT 287
Db 1233 CCYTACCCCTCCCTTAATAAAACAATCTTCT 1205

RESULT 50

ABL33538/C
ID ABL33538 standard; DNA; 17738 BP.

AC ABL33538;

XX 26-MAR-2002 (first entry)

XX Human immune system associated gene SEQ ID NO: 1511.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytotatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
ds.

XX Homo sapiens.

XX WO200200928-A2.

XX 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP007537.

XX 30-JUN-2000; 2000DE-01032529.

XX 01-SEP-2000; 2000DE-01043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful for
XX diagnosis and treatment of diseases associated with abnormal cytosine
XX methylation.

XX Claim 1; SEQ ID NO 1511; 32pp + Sequence Listing; German.

XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention

XX Sequence 17738 BP; 4696 A; 289 C; 3655 G; 9098 T; 0 U; 0 Other;

XX Query Match 12.1%; Score 36.2; DB 6; Length 17738;
XX Best Local Similarity 51.6%; Pred. No. 11;
XX Matches 83; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Qy 40 CTTGAATATTCTTCATTTGAAATAACAGCTAACTCCCAATTTGTACCTAACTAGGGGA 99
Db 9292 CTTATCTTATTTCAAATCTTAAAAAAAATCTTTCTTAATTTATACACAATACATA 9233

Qy 100 GTTTATCATGAGAAATGTTTAAATGCTTATTTTCAATAAGTAGTACCAGATGTTG 159
Db 9232 TAATCTTATCTATCAAAAAAAAATAAATAATTTAAAAAAAATTCAAAAAAATTTTA 9173

Qy 160 CAACATGCTGTTAATAAATACCCGACAAAACCTTCAATCACTATA 200
Db 9172 AAAAAACATTATATAAATTTATCTCAAAAACCTTAAAAAACTAAA 9132

Search completed: July 17, 2006, 21:54:57
Job time : 396 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 21:55:10 ; Search time 2892 Seconds
(without alignments)
5800.761 Million cell updates/sec

Title: SEQ1-96535C
Perfect score: 300
Sequence: 1 ggcattgacggctaaatgc.....atctgctgattaggaagat 300

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum Match 0%

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

EST.*

1: gb_est1.*

2: gb_est3.*

3: gb_est4.*

4: gb_est5.*

5: gb_est6.*

6: gb_hcc2.*

7: gb_est2.*

8: gb_est7.*

9: gb_est8.*

10: gb_est9.*

11: gb_gss1.*

12: gb_gss2.*

13: gb_gss3.*

14: gb_gss4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 141.4 | 47.1 | 715 | 13 | CW889980 |
| 2 | 88.2 | 29.4 | 183 | 13 | DU405174 |
| 3 | 45.8 | 15.3 | 684 | 2 | BJ924615 |
| 4 | 45.8 | 15.3 | 765 | 2 | BJ919954 |
| 5 | 43.6 | 14.5 | 411 | 8 | C0338585 |
| 6 | 43.6 | 14.5 | 514 | 8 | C0275038 |
| 7 | 43.6 | 14.5 | 545 | 8 | C0264525 |
| 8 | 43.6 | 14.5 | 604 | 8 | C0312735 |
| 9 | 43 | 14.3 | 818 | 11 | AZ123198 |
| 10 | 42 | 14.0 | 399 | 8 | C0185341 |
| 11 | 42 | 14.0 | 1933 | 6 | AK038018 |
| 12 | 41.4 | 13.8 | 684 | 11 | BZ281627 |
| 13 | 40.8 | 13.6 | 674 | 9 | DN248758 |
| 14 | 40.4 | 13.5 | 788 | 11 | BH463846 |
| 15 | 40.2 | 13.4 | 694 | 13 | CW144136 |
| 16 | 40.2 | 13.4 | 1172 | 12 | CC272514 |
| 17 | 39.6 | 13.2 | 543 | 5 | CD288289 |
| 18 | 39.6 | 13.2 | 712 | 5 | CJ361302 |
| 19 | 39.6 | 13.2 | 743 | 14 | CT404992 |

| | | | | | |
|----|------|------|------|----|----------|
| 20 | 39.6 | 13.2 | 745 | 5 | CJ381709 |
| 21 | 39.6 | 13.1 | 810 | 5 | CJ385948 |
| 22 | 39.6 | 13.1 | 539 | 13 | CZ012938 |
| 23 | 39.4 | 13.1 | 758 | 14 | DU930095 |
| 24 | 39.4 | 12.9 | 547 | 12 | CG804676 |
| 25 | 38.8 | 12.9 | 690 | 11 | BH929322 |
| 26 | 38.8 | 12.9 | 359 | 7 | BF336378 |
| 27 | 38.6 | 12.9 | 759 | 11 | BZ049137 |
| 28 | 38.4 | 12.8 | 535 | 14 | CR068462 |
| 29 | 38.4 | 12.8 | 683 | 13 | CL341203 |
| 30 | 38.4 | 12.8 | 779 | 9 | DN376396 |
| 31 | 38.4 | 12.8 | 746 | 9 | CK973570 |
| 32 | 38.2 | 12.7 | 756 | 8 | CV975922 |
| 33 | 38.2 | 12.7 | 759 | 11 | BH508774 |
| 34 | 38.2 | 12.7 | 808 | 13 | CZ764379 |
| 35 | 38.2 | 12.7 | 818 | 13 | DU208665 |
| 36 | 38.2 | 12.7 | 3792 | 11 | AF156725 |
| 37 | 38 | 12.7 | 559 | 11 | AO836871 |
| 38 | 38 | 12.7 | 1091 | 10 | DM602832 |
| 39 | 37.8 | 12.6 | 914 | 13 | CZ953670 |
| 40 | 37.8 | 12.6 | 950 | 14 | DU931319 |
| 41 | 37.6 | 12.5 | 684 | 4 | CB430880 |
| 42 | 37.6 | 12.5 | 714 | 11 | AO855958 |
| 43 | 37.6 | 12.5 | 837 | 11 | BH486214 |
| 44 | 37.6 | 12.5 | 991 | 3 | BU355289 |
| 45 | 37.4 | 12.5 | 373 | 10 | R22138 |
| 46 | 37.4 | 12.5 | 514 | 11 | AQ321454 |
| 47 | 37.4 | 12.5 | 563 | 11 | AQ627491 |
| 48 | 37.4 | 12.5 | 704 | 13 | CW685077 |
| 49 | 37.4 | 12.5 | 723 | 11 | BZ047105 |
| 50 | 37.4 | 12.5 | 797 | 11 | BZ479314 |
| 51 | 37.4 | 12.5 | 896 | 13 | CZ529782 |
| 52 | 37.4 | 12.5 | 940 | 13 | DU080282 |
| 53 | 37.2 | 12.4 | 416 | 8 | CO343761 |
| 54 | 37.2 | 12.4 | 486 | 7 | BE856330 |
| 55 | 37.2 | 12.4 | 521 | 1 | AV606314 |
| 56 | 37.2 | 12.4 | 551 | 1 | AV602756 |
| 57 | 37.2 | 12.4 | 647 | 4 | CB446831 |
| 58 | 37.2 | 12.4 | 660 | 4 | CB425397 |
| 59 | 37.2 | 12.4 | 672 | 8 | CN440384 |
| 60 | 37.2 | 12.4 | 683 | 4 | CB464131 |
| 61 | 37.2 | 12.4 | 693 | 8 | CV974819 |
| 62 | 37.2 | 12.4 | 700 | 8 | CO877917 |
| 63 | 37.2 | 12.4 | 788 | 9 | CK713951 |
| 64 | 37.2 | 12.4 | 795 | 9 | DN545081 |
| 65 | 37.2 | 12.4 | 830 | 9 | DN543711 |
| 66 | 37.2 | 12.4 | 1080 | 1 | AJ925742 |
| 67 | 37.2 | 12.4 | 1253 | 10 | DV779655 |
| 68 | 37 | 12.3 | 366 | 5 | CK779724 |
| 69 | 37 | 12.3 | 527 | 4 | BX490203 |
| 70 | 37 | 12.3 | 613 | 4 | BX641441 |
| 71 | 37 | 12.3 | 618 | 11 | AQ244513 |
| 72 | 37 | 12.3 | 621 | 8 | CN762630 |
| 73 | 37 | 12.3 | 628 | 13 | CW144135 |
| 74 | 37 | 12.3 | 683 | 4 | BX504981 |
| 75 | 37 | 12.3 | 707 | 4 | EX509011 |
| 76 | 37 | 12.3 | 733 | 14 | AG037123 |
| 77 | 37 | 12.3 | 778 | 13 | CZ465308 |
| 78 | 37 | 12.3 | 805 | 13 | DU209109 |
| 79 | 37 | 12.3 | 809 | 14 | CT185225 |
| 80 | 37 | 12.3 | 884 | 14 | DU888537 |
| 81 | 37 | 12.3 | 905 | 3 | BQ948933 |
| 82 | 37 | 12.3 | 945 | 14 | CNS04D0K |
| 83 | 36.8 | 12.3 | 301 | 2 | BG461576 |
| 84 | 36.8 | 12.3 | 307 | 2 | BG461289 |
| 85 | 36.8 | 12.3 | 368 | 8 | CN360568 |
| 86 | 36.8 | 12.3 | 498 | 3 | BQ024056 |
| 87 | 36.8 | 12.3 | 558 | 8 | CV221846 |
| 88 | 36.8 | 12.3 | 566 | 3 | BU752970 |
| 89 | 36.8 | 12.3 | 632 | 7 | BE631244 |
| 90 | 36.8 | 12.3 | 642 | 3 | BN985293 |
| 91 | 36.8 | 12.3 | 676 | 3 | BN683514 |
| 92 | 36.8 | 12.3 | 680 | 4 | CA056323 |

| | |
|----------|------------|
| CJ381709 | CJ381709 |
| CJ385948 | CJ385948 |
| CZ012938 | CH240_507 |
| DU930095 | 349135 To |
| CG804676 | 1118052E0 |
| BH929322 | odi02408. |
| BF336378 | CM3-CT051 |
| BZ049137 | jnr21908. |
| CR068462 | Forward s |
| CL341203 | RPC144_26. |
| DN376396 | LIB38529 |
| CK973570 | JGI_CAAAP8 |
| CV975922 | UMC-bcl_0 |
| BH508774 | BOGQB34TF |
| CZ764379 | OC_Ba011 |
| DU208665 | 109857424 |
| AF156725 | AF156725 |
| AO836871 | HS_5463_A |
| DM602832 | CGX162-D0 |
| CZ953670 | 265847 To |
| DU931319 | 353827 To |
| CB430880 | 606805 MA |
| AO855958 | CpG1511A |
| BH486214 | BOGKS58TR |
| BU355289 | 603474334 |
| R22138 | yh25D10.s1 |
| AQ321454 | RPC111-10 |
| AQ627491 | CITBI-E1- |
| CW685077 | OG_BBA004 |
| BZ047105 | 1kF8806 |
| BZ479314 | BONHM39TF |
| CZ529782 | SRAA-aac6 |
| DU080282 | 55668 Tom |
| CO343761 | EP18004_3 |
| BE856330 | 7f93a05.x |
| AV606314 | AV602756 |
| CB446831 | 700566 MA |
| CB425397 | 600378 MA |
| CN440384 | BE04022A1 |
| CB464131 | 725316 MA |
| CV974819 | UMC-bcl_0 |
| CO877917 | Bovgen_06 |
| CK713951 | RTPO1_15 |
| DN545081 | 1394251 M |
| DN543711 | 1392585 M |
| AJ925742 | AJ925742 |
| DV779655 | Hw_Fat_72 |
| CK779724 | UMC-bend |
| BX490203 | DKF2p686G |
| BX641441 | DKF2p686E |
| AQ244513 | HS_2050_A |
| CN762630 | ID0AAACH |
| CW144135 | 104_536_1 |
| BX504981 | DKF2p686K |
| EX509011 | DKF2p686O |
| AG037123 | Pan_trog1 |
| CZ465308 | MC756119 |
| DU209109 | 109857426 |
| CT185225 | Sus scrof |
| DU888537 | 390210 To |
| BQ948933 | AGENCOURT |
| AL285149 | Tetraodon |
| BG461576 | RST4441 |
| BG461289 | RST44074 |
| CN360568 | 170005999 |
| BQ024056 | UI-1-BB1p |
| CV221846 | le_mx0_53 |
| BU752970 | UI-1-BB1- |
| BE631244 | BB631244 |
| BN985293 | UI-CP-EC1 |
| BN683514 | UI-E-EO1- |
| CA056323 | ssalrgb50 |

| | | | | | | |
|-------|------|------|------|----|----------|------------|
| c 93 | 36.8 | 12.3 | 727 | 13 | CL832859 | OR_CBa005 |
| c 94 | 36.8 | 12.3 | 732 | 9 | DN637524 | UMC_bcl_0 |
| c 95 | 36.8 | 12.3 | 742 | 11 | BZ074426 | Ljr9A009. |
| c 96 | 36.8 | 12.3 | 777 | 13 | CL847336 | OR_CBa007 |
| c 97 | 36.8 | 12.3 | 831 | 13 | CL762851 | OR_BBa013 |
| c 98 | 36.8 | 12.3 | 837 | 13 | CL615241 | OA_ABa015 |
| c 99 | 36.8 | 12.3 | 897 | 2 | BG402233 | 602465796 |
| c 100 | 36.8 | 12.3 | 901 | 14 | DU946208 | 375232 To |
| c 101 | 36.8 | 12.3 | 941 | 13 | CW940128 | TGB20.4 G |
| c 102 | 36.8 | 12.3 | 959 | 3 | BU503904 | AGENCOURT |
| c 103 | 36.8 | 12.3 | 983 | 12 | CC235815 | CH261-139 |
| c 104 | 36.8 | 12.3 | 2457 | 6 | AK037941 | Mus muscu |
| c 105 | 36.8 | 12.2 | 288 | 2 | BG460117 | RST42571 |
| c 106 | 36.6 | 12.2 | 333 | 13 | CL868795 | abe52a03. |
| c 107 | 36.6 | 12.2 | 638 | 11 | BH743817 | gt27e09.b |
| c 108 | 36.6 | 12.2 | 710 | 11 | BH999124 | oeh96d10. |
| c 109 | 36.6 | 12.2 | 713 | 12 | CE509108 | tigr-gss- |
| c 110 | 36.6 | 12.2 | 739 | 11 | BZ071205 | lkg56g07. |
| c 111 | 36.6 | 12.2 | 760 | 14 | AG591277 | Mus muscu |
| c 112 | 36.6 | 12.2 | 793 | 11 | BZ158667 | CH230-289 |
| c 113 | 36.6 | 12.2 | 798 | 11 | BZ476138 | BONIL13TR |
| c 114 | 36.6 | 12.2 | 842 | 14 | DU979127 | KBrHI0651 |
| c 115 | 36.6 | 12.2 | 871 | 14 | CT017013 | KBrHI28P2 |
| c 116 | 36.6 | 12.2 | 897 | 14 | DU919305 | 338842 To |
| c 117 | 36.6 | 12.2 | 955 | 14 | DU949088 | 235939 To |
| c 118 | 36.6 | 12.2 | 1007 | 11 | AZ139311 | SP_0179_B |
| c 119 | 36.6 | 12.2 | 1186 | 8 | CN641719 | ILLUMIGEN |
| c 120 | 36.4 | 12.1 | 320 | 10 | H56086 | Vq97g05.s1 |
| c 121 | 36.4 | 12.1 | 392 | 10 | R20585 | Vj05b10.r1 |
| c 122 | 36.4 | 12.1 | 396 | 10 | R81837 | |
| c 123 | 36.4 | 12.1 | 420 | 13 | CW648590 | OA_ABa018 |
| c 124 | 36.4 | 12.1 | 576 | 5 | CK772214 | 960730 MA |
| c 125 | 36.4 | 12.1 | 596 | 12 | CG680618 | 74J19_03T |
| c 126 | 36.4 | 12.1 | 604 | 12 | CE152910 | tigr-gss- |
| c 127 | 36.4 | 12.1 | 619 | 12 | CE412104 | tigr-gss- |
| c 128 | 36.4 | 12.1 | 677 | 13 | CW400942 | fsbb001f0 |
| c 129 | 36.4 | 12.1 | 684 | 13 | CW400941 | fsbb001f0 |
| c 130 | 36.4 | 12.1 | 713 | 10 | DT468106 | GH_CHX2IG |
| c 131 | 36.4 | 12.1 | 741 | 13 | DU295356 | 109844801 |
| c 132 | 36.4 | 12.1 | 779 | 11 | AQ312092 | RPC111-93 |
| c 133 | 36.4 | 12.1 | 783 | 13 | CL913041 | OA_ABa001 |
| c 134 | 36.4 | 12.1 | 843 | 11 | AZ668782 | ENTL152TF |
| c 135 | 36.4 | 12.1 | 879 | 14 | CT319003 | Sus scrofa |
| c 136 | 36.4 | 12.1 | 1007 | 14 | DU762740 | ASNG537.b |
| c 137 | 36.2 | 12.1 | 357 | 2 | BM501879 | ih65a07.y |
| c 138 | 36.2 | 12.1 | 401 | 3 | BM730934 | ih65a07.x |
| c 139 | 36.2 | 12.1 | 457 | 7 | AW334276 | S32P11 AG |
| c 140 | 36.2 | 12.1 | 592 | 5 | CJ392037 | CJ392037 |
| c 141 | 36.2 | 12.1 | 718 | 7 | AW334928 | S40H11 AG |
| c 142 | 36.2 | 12.1 | 721 | 9 | DN245644 | ACAE-aaa2 |
| c 143 | 36.2 | 12.1 | 722 | 7 | AW334756 | S37F6 AGS |
| c 144 | 36.2 | 12.1 | 755 | 13 | CW503920 | OP_Ba001 |
| c 145 | 36.2 | 12.1 | 771 | 13 | CL709159 | OR_BBa003 |
| c 146 | 36.2 | 12.1 | 772 | 13 | CW830602 | OP_Ba009 |
| c 147 | 36.2 | 12.1 | 801 | 14 | AG353786 | Mus muscu |
| c 148 | 36.2 | 12.1 | 828 | 11 | AZ533640 | ENTCN67TR |
| c 149 | 36.2 | 12.1 | 900 | 11 | BH134888 | ENTO154TR |
| c 150 | 36.2 | 12.1 | 914 | 11 | AZ684348 | ENTME12TR |

ALIGNMENTS

| | | | | | |
|------------|--------------------------|-------------|------------|---------------|-----------------|
| RESULT 1 | CW889980 | linear | DNA | 715 bp | GSS 08-DEC-2004 |
| LOCUS | RPC142_112H1.TV | RPCI-42 | Bos taurus | genomic clone | RPC142_112H1, |
| DEFINITION | genomic survey sequence. | | | | |
| ACCESSION | CW889980 | | | | |
| VERSION | CW889980.1 | GI:56429161 | | | |
| KEYWORDS | GSS. | | | | |
| SOURCE | Bos taurus | (cattle) | | | |
| ORGANISM | Bos taurus | | | | |

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

1 (bases 1 to 715)

Larkin,D.M., Donthu,K., LeDuc,R., Ryan,K., Liu,L., de Jong,P.J. and Lewin,H.A.

End sequencing of Holstein BAC library RPCI-42

Unpublished (2004)

Other GSSs: RPC142_112H1.TJ

Contact: Harris Lewin

Department of Animal Sciences

University of Illinois at Urbana Champaign

1201 W. Gregory Dr., Urbana, IL 61801, USA

Tel: 217 333 5998

Fax: 217 244 5617

Email: h-lewin@uiuc.edu

Clones are derived from the bovine BAC library RPCI-42 (http://bacpac.choi.org/mbovine42.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.choi.org). Clones may be purchased from BACPAC Resources (http://www.choi.org/bacpac/ordering_information.htm). Funding was provided in part by grant no. AG2004-34480-14417 from USDA-CSREES (Livestock Genome Sequencing Initiative) and AG58-5438-2-313 from USDA-ARS

Plate: 112 Row: H Column: 1

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..715

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="Holstein"

/db_xref="taxon:9913"

/clone="RPC142_112H1"

/sex="male"

/cell_type="blood"

/clone_lib="RPCI-42"

/notes="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; RPCI-42 Bovine BAC library (Male) produced by Pieter de Jong"

Query Match 47.1%; Score 141.4; DB 13; Length 715;

Best Local Similarity 77.6%; Pred. No. 5.2e-26;

Matches 184; Conservative 0; Mismatches 51; Indels 2; Gaps 1;

QY 64 AAACAGCTAACTCCCAATTTGTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAA 123

Db 263 AAACAGCCAGCTCCCAATTTGTCCATTAACCTAGCGAAATCCATCATGAAGAAATG--TAAA 320

QY 124 TGCTTAATTTTCAAATAAGTAAGTAACAGAGTGTTCGAAATGCTGTGTTAAATACCCGAC 183

Db 321 AGCTTATTTTAAATAAATAAAGTCACCAGAGTATTGCAACATCTCTGTTAAACAGCAAAAC 380

QY 184 AAACCTTCAATCACTATAGCTAGTAGAGTGCAATCTTCGCAAGATCCAGAGTAACCCAGT 243

Db 381 AAATTCACCTCATTACAGCTGTAGTAGACTGCGTCTCTGTAGAAATTTTCAGAGTAACGAT 440

QY 244 ATTTTGGAAATGCAATGTTGAAACCGACATACATAATTTATCTGCTGATTAGGAAGTAT 300

Db 441 ATTTTGGAAATGCAATGTTGAAACCGACATACATAATTTATCTGCTGATTAGGAAGTAT 497

RESULT 2

DU405174

LOCUS 1098448345721 CHORI-243 Ovis aries genomic clone CH243-193BB,

DEFINITION genomic survey sequence.

ACCESSION DU405174

VERSION DU405174.1 GI:77135304

KEYWORDS GSS.

SOURCE Ovis aries (sheep)

ORGANISM Ovis aries


```

/dev_stage="adult"
/clone_lib="MF01ALA cdna"

Query Match      15.3%; Score 45.8; DB 2; Length 765;
Best Local Similarity 53.0%; Pred. No. 0.37;
Matches 98; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 36 AAGACTTGAATATCTTCAATTTGAATAAACAAGCTAACTCCCAAAATGTACCTAACTAG 95
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 325 AAGAAGCTAAATCCCTATCTCTGAATAAATAAAGAGCTGTATATATGTTATGTATAG 266
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 96 GGGAGTTATCATGACGAATATGTTAAATGCTTAATTTTCAATAAAGAAGTAACCAAGT 155
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 265 ATGAATAGACTTTTGAAAAAATGTTAAAAAATGACAAGTTGGGAAGTGTTATATCAAGAAT 206
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 156 GTTGCAACATGCTGTTTAAATAACCCGACAACTCAATCACTACTAGCTGAGTAGAGTGC 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 205 TGAGGAATAACTCTCAAAAAACATTCCACTTTAATAATAATTTGTTAATAATTC 146
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 216 ATTCT 220
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 145 AATTT 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
CO338585/c
LOCUS      CO338585      411 bp      mRNA      linear      EST 05-OCT-2004
DEFINITION EN17132.5prime Exelixis FlyTag MN08 BlueScript Drosophila
            melanogaster cdna clone EN17132 5, mRNA sequence.
ACCESSION  CO338585
VERSION     CO338585.1 GI:49398880
KEYWORDS
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 411)
AUTHORS    Nakanishi,M., Muzong,C., Peterson,E., Laufer,A., Leung,W., Platt,D.
            and Swimmer,C.
TITLE      Exelixis FlyTag EST Project MN08 Library
JOURNAL    Unpublished (2004)
COMMENT    Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: EN.171 row: C column: 8
            High quality sequence stop: 410.
            Location/Qualifiers
FEATURES             source
     source           1..411
                     /organism="Drosophila melanogaster"
                     /mol_type="mRNA"
                     /db_xref="taxon:7227"
                     /clone="EN17132"
                     /clone_lib="Exelixis FlyTag MN08 pCDNA-SK+"
                     /note="Organ: mixed stage embryos, imaginal disks, and
                     adult heads; Vector: pCDNA-SK+; Site 1: NotI; Site 2:
                     XhoI; Random primed, normalized library from mixed stage
                     embryos, imaginal disks, and adult heads."
ORIGIN
Query Match      14.5%; Score 43.6; DB 8; Length 411;
Best Local Similarity 58.5%; Pred. No. 1.3;
Matches 76; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 29 TACGATGAAGACTTGAATATCTTCAATTTGAATAAACAAGCTAACTCCCAAAATGTACC 88
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 TATAATTAAGACTTAAAAAATATTTTTTTTTTAAAAAATAATTTATTTTAAATTAATACT 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 89 TAACTAGGGGAGTTTATCATGAAGAATAATGTTTAAATGCTTTAAATTTCAAATAAGAGTAA 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111 TAATAAATTATTTTATATAAATATTTTATATAAATAAATTTATTTTAAATAATTAATAA 52
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 149 CCAGAGTGTT 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 51 AAATATTTTT 42
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
CO264525/c
LOCUS      CO264525      545 bp      mRNA      linear      EST 23-JUN-2004
DEFINITION EK139457.5prime Exelixis FlyTag CK01 pCDNA-SK+ Drosophila
            melanogaster cdna clone EK139457 5, mRNA sequence.
ACCESSION  CO264525
VERSION     CO264525.1 GI:49150110
KEYWORDS
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 514)
AUTHORS    Kopczynski,C., Platt,D., Campbell,J., Muzong,C., Laufer,A.,
            Peterson,E. and Swimmer,C.
TITLE      Exelixis FlyTag EST Project CK01 Library
JOURNAL    Unpublished (2004)
COMMENT    Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: EK.1162 row: E column: 9
            High quality sequence stop: 407.
            Location/Qualifiers
FEATURES             source
     source           1..514
                     /organism="Drosophila melanogaster"
                     /mol_type="mRNA"
                     /db_xref="taxon:7227"
                     /clone="EK116257"
                     /clone_lib="Exelixis FlyTag CK01 pCDNA-SK+"
                     /note="Organ: mixed stage embryos, imaginal disks, and
                     adult heads; Vector: pCDNA-SK+; Site 1: NotI; Site 2:
                     XhoI; Random primed, normalized library from mixed stage
                     embryos, imaginal disks, and adult heads."
ORIGIN
Query Match      14.5%; Score 43.6; DB 8; Length 514;
Best Local Similarity 58.5%; Pred. No. 1.3;
Matches 76; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 29 TACGATGAAGACTTGAATATCTTCAATTTGAATAAACAAGCTAACTCCCAAAATGTACC 88
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 TATAATTAAGACTTAAAAAATATTTTTTTTTTAAAAAATAATTTATTTTAAATTAATACT 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 89 TAACTAGGGGAGTTTATCATGAAGAATAATGTTTAAATGCTTTAAATTTCAAATAAGAGTAA 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111 TAATAAATTATTTTATATAAATATTTTATATAAATAAATTTATTTTAAATAATTAATAA 52
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 149 CCAGAGTGTT 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 51 AAATATTTTT 42
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
CO264525/c
LOCUS      CO264525      545 bp      mRNA      linear      EST 23-JUN-2004
DEFINITION EK139457.5prime Exelixis FlyTag CK01 pCDNA-SK+ Drosophila
            melanogaster cdna clone EK139457 5, mRNA sequence.
ACCESSION  CO264525
VERSION     CO264525.1 GI:49150110
KEYWORDS
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 514)
AUTHORS    Kopczynski,C., Platt,D., Campbell,J., Muzong,C., Laufer,A.,
            Peterson,E. and Swimmer,C.
TITLE      Exelixis FlyTag EST Project CK01 Library
JOURNAL    Unpublished (2004)
COMMENT    Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: EK.1162 row: E column: 9
            High quality sequence stop: 407.
            Location/Qualifiers
FEATURES             source
     source           1..514
                     /organism="Drosophila melanogaster"
                     /mol_type="mRNA"
                     /db_xref="taxon:7227"
                     /clone="EK116257"
                     /clone_lib="Exelixis FlyTag CK01 pCDNA-SK+"
                     /note="Organ: mixed stage embryos, imaginal disks, and
                     adult heads; Vector: pCDNA-SK+; Site 1: NotI; Site 2:
                     XhoI; Random primed, normalized library from mixed stage
                     embryos, imaginal disks, and adult heads."
ORIGIN
Query Match      14.5%; Score 43.6; DB 8; Length 411;
Best Local Similarity 58.5%; Pred. No. 1.3;
Matches 76; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 29 TACGATGAAGACTTGAATATCTTCAATTTGAATAAACAAGCTAACTCCCAAAATGTACC 88
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 304 TATAATTAAGACTTAAAAAATATTTTTTTTTTAAAAAATAATGATTTTAAATTAATACT 245
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 545)
AUTHORS Kopczyński, C., Platt, D., Campbell, J., Muzong, C., Laufer, A.,
Peterson, E. and Swimmer, C.
TITLE Exelixis FlyTag EST Project CK01 Library
JOURNAL Unpublished (2004)
COMMENT Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Based upon one or more reads of this clone where vector sequence
was present at both ends, this clone has been determined to contain
contain a cDNA insert on the order of 600-1000 bases.
Plate: EK.1394 row: E column: 9
High quality sequence stop: 541.
Location/Qualifiers
1. 545
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="EK139457"
/clone_lib="Exelixis FlyTag CK01 pCDNA-SK+"
/notes="Organ: mixed stage embryos, imaginal disks, and
adult heads; Vector: pCDNA-SK+; Site 1: NotI; Site 2:
XhoI; Random primed, normalized library from mixed stage
embryos, imaginal disks, and adult heads."

ORIGIN
Query Match 14.5%; Score 43.6; DB 8; Length 545;
Best Local Similarity 58.5%; Pred. NO. 1.3;
Matches 76; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 29 TACGATGAAGACTTGAATATTTCTTCATTGCAATTAACAGCTAATCCCAATTGTACC 88
DB 328 TATATTAACACTTAAATAATTTTCTTTTAAATAAAATTTATTTATTAATTTACT 269
QY 89 TAATAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAATAAGAGTAA 148
DB 268 TAATAAAGCTATTTTATTAATAAATTTTATTAATAAATTTTATTAATAAATTAATA 209
QY 149 CCAGAGTGT 158
DB 208 AAATATTTT 199

RESULT 8
LOCUS CO312735/6 604 bp mRNA linear EST 28-JUN-2004
DEFINITION EK229834.5prime Exelixis FlyTag CK01 pCDNA-SK+ Drosophila
melanogaster cDNA clone EK229834 5, mRNA sequence.
ACCESSION CO312735
VERSION CO312735.1 GI:49368436
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 604)
AUTHORS Kopczyński, C., Platt, D., Campbell, J., Muzong, C., Laufer, A.,
Peterson, E. and Swimmer, C.
TITLE Exelixis FlyTag EST Project CK01 Library
JOURNAL Unpublished (2004)
COMMENT Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Based upon one or more reads of this clone where vector sequence
was present at both ends, this clone has been determined to contain
contain a cDNA insert on the order of 600-1000 bases.
Plate: EK.1394 row: E column: 9
High quality sequence stop: 541.
Location/Qualifiers
1. 545
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="EK139457"
/clone_lib="Exelixis FlyTag CK01 pCDNA-SK+"
/notes="Organ: mixed stage embryos, imaginal disks, and
adult heads; Vector: pCDNA-SK+; Site 1: NotI; Site 2:
XhoI; Random primed, normalized library from mixed stage
embryos, imaginal disks, and adult heads."

ORIGIN
Query Match 14.5%; Score 43.6; DB 8; Length 604;
Best Local Similarity 58.5%; Pred. NO. 1.3;
Matches 76; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 29 TACGATGAAGACTTGAATATTTCTTCATTGCAATTAACAGCTAATCCCAATTGTACC 88
DB 329 TATAATTAAGACTTAAATAATTTTCTTTTAAATAAAATTTATTTATTAATTTACT 270
QY 89 TAATAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAATAAGAGTAA 148
DB 269 TAATAAAGCTATTTTATTAATAAATTTTATTAATAAATTTTATTAATAAATTAATA 210
QY 149 CCAGAGTGT 158
DB 209 AAATATTTT 200

RESULT 9
LOCUS AZ123198 818 bp DNA linear GSS 12-MAY-2000
DEFINITION RPCI-23-22D9.TJ RPCI-23 Mus musculus genomic clone RPCI-23-22D9,
genomic survey sequence.
ACCESSION AZ123198
VERSION AZ123198.1 GI:7791733
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 818)
AUTHORS Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S.,
Akinret, B., Levins, M., McGann, S., Teegaye, G., Geer, K., Krol, M., de
Jong, P. and Fraser, C.M.
TITLE Mouse BAC End Sequences from Library RPCI-23
JOURNAL Unpublished (1999)
COMMENT Other GSSs: RPCI-23-22D9.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac/bac\_ends/mouse/bac\_end\_intro.html
Plate: 22 row: D column: 9
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. 818

FEATURES
source

```

```

/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPC1-23-22D9"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPC1-23"
/notes="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
ECORI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
ECORI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

```

ORIGIN

```

Query Match      14.3%; Score 43; DB 11; Length 818;
Best Local Similarity 56.8%; Pred. No. 2;
Matches 79; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 111 AGAAATGTTTAAATGCTTAATTTTCAATAAGAGTAACAGAGTGTTCGAACATGCTGT 170
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 301 ACAATGATCATGTGACTAATTATATTATTAAGGAAAAAGCATGCTGAAGGAGTA 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 171 TAAATAACCCGACAACTCAATCACTATAGCTGTAGTAGTGCAATTCGCAAGGATCC 230
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 361 AAGTTAGCTTAAGAAATTTTAAAAATAATAGATTGAGTAGATGATGAGCAAGGACAC 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 231 CAGAGTAACCAATTTTGT 249
    ||||| ||||| ||||| |||||
DB 421 AAGACTATCTCTTAATTG 439
    ||||| ||||| ||||| |||||

```

RESULT 10

```

C0185341/c
LOCUS      399 bp mRNA linear EST 21-JUN-2004
DEFINITION Exelixis FlyTag CK01 pCDNA-SK+ Drosophila
            melanogaster cDNA clone EK008250 5, mRNA sequence.
ACCESSION  C0185341
VERSION     C0185341.1 GI:48996516
KEYWORDS   EST.
SOURCE     Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 399)
            Kopczynski, C., Platt, D., Campbell, J., Muzong, C., Laufer, A.,
            Peterson, E., and Swimmer, C.
            Exelixis FlyTag EST Project CK01 Library
            Unpublished (2004)
            Contact: Stapleton, M.
            BGGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST\_est@fruitfly.berkeley.edu
            Plate: EK 82 row: E column: 2
            High quality sequence stop: 374.
            Location/Qualifiers
                1..399
                /organism="Drosophila melanogaster"
                /mol_type="mRNA"
                /db_xref="taxon:7227"
                /clone="EK008250"
                /clone_lib="Exelixis FlyTag CK01 pCDNA-SK+"
                /notes="Organ: mixed stage embryos, imaginal disks, and
                adult heads; Vector: pCDNA-SK+; Site 1: NotI; Site 2:
                XhoI; Random primed, normalized library from mixed stage
                embryos, imaginal disks, and adult heads."

```

FEATURES

```

source
1..399
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="EK008250"
/clone_lib="Exelixis FlyTag CK01 pCDNA-SK+"
/notes="Organ: mixed stage embryos, imaginal disks, and
adult heads; Vector: pCDNA-SK+; Site 1: NotI; Site 2:
XhoI; Random primed, normalized library from mixed stage
embryos, imaginal disks, and adult heads."

```

ORIGIN

```

Query Match      14.0%; Score 42; DB 8; Length 399;
Best Local Similarity 57.7%; Pred. No. 3.3;
Matches 75; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 29 TACGATGAAGACTTGAAATATTCTTCATTTTGAATAAAGCAAGTCCCAAAATTGTACC 88
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 388 TATAATAAACTTAAAAAATATTTTTTTTTTAAAAAAAATATTTATTATTAAATTATAT 329
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 89 TAACTAGGGGAGTTTATCATGAAGAAATGTTTAAATGCTTTCAATAAGAAGTAA 148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 328 TAATAAAATATTTTATAATAAATATTTTATAAATAAATAATTTTAAAAATAATAATA 269
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 149 CCAGAGTGTT 158
    ||||| ||||| ||||| |||||
DB 268 AAATATTTTT 259
    ||||| ||||| ||||| |||||

```

RESULT 11

```

AK038018
LOCUS      1933 bp mRNA linear HTC 02-SEP-2005
DEFINITION Mus musculus 16 days neonate thymus cDNA, RIKEN full-length
            enriched library, clone:A130072A19 product:unclassifiable, full
            insert sequence.
ACCESSION  AK038018
VERSION     AK038018.1 GI:26086232
KEYWORDS   HTC; CAP trapper.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE  1
            Carninci, P. and Hayashizaki, Y.
            High-efficiency full-length cDNA cloning
            Meth. Enzymol. 303, 19-44 (1999)
            10349636
REFERENCE  2
            Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
            Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
            Genome Res. 10 (10), 1617-1630 (2000)
            11042159
REFERENCE  3
            Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
            Konno, H., Akiyama, J., Nishi, K., Kitsuimui, T., Tashiro, H., Itoh, M.,
            Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
            Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
            Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
            Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
            Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
            RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
            Genome Res. 10 (11), 1757-1771 (2000)
            11076861
REFERENCE  4
            The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
            Functional annotation of a full-length mouse cDNA collection
            Nature 409, 685-690 (2001)
REFERENCE  5
            The FANTOM Consortium, the RIKEN Genome Exploration Research Group
            Phase I and II Team.
            Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
            Nature 420, 563-573 (2002)
REFERENCE  6
            RIKEN Genome Exploration Research Group, Genome Science Group
            (Genome Network Core Team) and the FANTOM Consortium.
            Antisense Transcription in the Mammalian Transcriptome
            Science 309, 1564-1566 (2005)
REFERENCE  7
            The FANTOM Consortium, Riken Genome Exploration Research Group and

```



```

VERSION      DN248758.2  GI:68340052
KEYWORDS
SOURCE       Hydra magnipapillata
ORGANISM     Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
              Hydridae; Hydra.
REFERENCE    1 (bases 1 to 674)
AUTHORS      Bode,H., Blumberg,B., Steele,R., Wigge,P., Gee,L., Nguyen,Q.,
              Martinez,D., Kibler,D., Hampson,S., Clifton,S., Pape,D., Marra,M.,
              Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B., Bowers,Y.,
              Gibbons,M., Ritter,E., Bennett,J., Ronko,I., Tsagareishvili,R.,
              Maguire,L., Kennedy,S., Waterston,R. and Wilson,R.
              WashU Hydra EST Project
TITLE        Unpublished (2002)
JOURNAL
COMMENT      On Mar 1, 2005 this sequence version replaced gi:60411218.
              Contact: Hans Bode
              WashU Hydra EST Project
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              Library material provided by Hans Bode & Dirk Lindgens, Univ. of
              Calif., Irvine Library constructed by Dirk Lindgens Univ. of
              Calif., Irvine Library sequenced by Washington University Genome
              Sequencing Center For information on obtaining a clone please
              contact: Hans Bode (hrobe@uci.edu)
              original OR value of 876 was extended to value 877 (,)
              This trace has been recalled with phred
              original value before phred recall for SL was 91
              original value before phred recall for SR was 879
              Seq primer: -40RP from Gibco.
FEATURES     source
              Location/Qualifiers
                1..674
                  /organism="Hydra magnipapillata"
                  /mol_type="mRNA"
                  /strain="105"
                  /db_xref="taxon:6085"
                  /lab_host="DH10B"
                  /clone_lib="Hydra EST UCI 5"
                  /notes="Vector: pSPORT1; Site 1: Not I; Site 2: Sal I;
                    a.1st strand cDNA was primed with a Not I primer-adaptor
                    (5' - PGACTAGTTCAGACGACGCGCCGCC(T)15-3')
                    b.Double-stranded cDNA was ligated to Sal I adaptor,
                    digested with Not I and cloned into the pSPORT 1-vector
                    pre-cut with Not I and Sal I. c.The ligation mix was
                    transformed into DH10B cells. d.The cells were grown in
                    SOC = 5% yeast, 20g tryptone, 0.5 g NaCl, 10 mM MgSO4, 10
                    mM MgCl, 0.2% glucose/liter, (no antibiotic). e.DMSO was
                    added to a final conc. of 10% as a cryoprotectant and
                    frozen f.The titre before freezing was determined as
                    ~2400/100 ul. Assuming a 10% loss upon freezing, the titre
                    is probably ~2100/ 100 ul. g.9 tubes each containing ~
                    2100 clones/100 ul (= total of ~19,000) are enclosed.
                    h.The frequency of vectors containing inserts is 96% as
                    determined by digestion check after picking 24 clones,
                    miniprep and subsequent digestion with Not I and Sal I.
                    i.A low level of 32p was used in the cDNA synthesis
                    procedure. The level measured by holding a Geiger Counter
                    next to a tube was background."
ORIGIN
Query Match      13.6%; Score 40.8; DB 9; Length 674;
Best Local Similarity 56.8%; Pred. No. 7.2;
Matches 75; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 56 TTTGAATAAACAGCTAACCTCCAAAATTGTACCTAACTAGGGAGGTTCATCATGAAGAAA 115
Db 418 TTTTAGACAACGATTTCTTCCAAAGTTTCAAAATTAGCTATGAGTGATGATCATAAAGACA 477
QY 116 TGTTTAATGCTTAATTTTCAAAATAAGAAGTAACACGAGTGTTCACATGCTCTTAAAT 175
Db 478 TTTTTCGATTAAACACTTTCGAATCTGGGGTTGCAAGACTGATACGTACTGCTGTAAAG 537

QY 176 AACCCGACAAAC 187
Db 538 CATTCATAAAC 549

RESULT 14
BH463846/c
LOCUS      BH463846
DEFINITION BOGVW38TF BOGV Brassica oleracea genomic clone BOGVW38, genomic
              survey sequence.
ACCESSION  BH463846
VERSION    BH463845.1 GI:17658643
KEYWORDS   GSS.
SOURCE     Brassica oleracea
ORGANISM   Brassica oleracea
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
              rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE  1 (bases 1 to 788)
AUTHORS    Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
              Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
              Whole genome shotgun sequencing of Brassica oleracea and its
              application to gene discovery and annotation in Arabidopsis
              Genome Res. 15 (4), 487-495 (2005)
JOURNAL    Genome Res. 15 (4), 487-495 (2005)
PUBMED     15805490
COMMENT     Other GSSs: BOGVW38TR
              Contact: Chris Town
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA.
              Tel: 301-838-3523
              Fax: 301-838-0208
              Email: cdtown@tigr.org
              DNA is from a doubled haploid provided by Tom Osborn.
              Seq primer: TF
              Class: sheared ends.
FEATURES     Location/Qualifiers
              1..788
                /organism="Brassica oleracea"
                /mol_type="genomic DNA"
                /strain="TO1000DH3"
                /db_xref="taxon:3712"
                /clone="BOGVW38"
                /notes="BOGV"
                /notes="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared
              genomic DNA inserted into PHOS1 using BstXI linkers"
ORIGIN
Query Match      13.5%; Score 40.4; DB 11; Length 788;
Best Local Similarity 47.3%; Pred. No. 9.4;
Matches 122; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 23 AGTGACTACGATGAAGACTTGAATAATTCTTCAATTTGAAATAAACAGCTAACTCCCAAT 82
Db 740 ACTGATTTTCGATTCAGATATATATTCATATATAGTTCATATTTGTTTGTGAA 681
QY 83 TGTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAAATAAG 142
Db 680 AGTACTTAATTTTATTTTCTATAAAAAATTTGTAAACAAATATTATTATTTATAATTT 621
QY 143 AAGTAACGAGAGTGTTCACATGCTGTTAAATAACCCGACCAACTTCATCACTATAGC 202
Db 620 TAGTACAAATTCGTATTAAAAATTTATGCAATTTAAAGCTACCAATCGATATTTTCAAAAC 561
QY 203 TGTAGTAGAGTGCATTCCTGCAAGGATCCCGAGATTAACAGTAACTTTTGGAAATGCAATGTT 262
Db 560 TGAATAATAATTTTTTTTTTATAAAATTTGTAATATGACCATGATTTTGTATGCGCTGAC 501
QY 263 GAACCGACCACTACTAATT 280
Db 500 CTGCCAACGTTCTGTGAGT 483

```

```

RESULT 15
CWI44136/c
LOCUS
DEFINITION
104 536_11137783.148 34960.028 Sorghum methylation filtered library
(LibID: 104) Sorghum bicolor genomic clone 11137783, genomic survey
sequence.
ACCESSION
CWI44136
VERSION
CWI44136.1 GI:54836683
SOURCE
Sorghum bicolor (sorghum)
ORGANISM
Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 694)
REFERENCE
AUTHORS
Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D.,
Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K.,
McMenamy, J., Smith, M., Holsman, H., Roe, B.A., Wiley, G., Korf, I.F.,
Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddloh, J.A. and
Martensen, R.A.
TITLE
Sorghum genome sequencing by methylation filtration
JOURNAL
PLOS Biol. 3 (1), e13 (2005)
PUBMED
15660154
COMMENT
Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: 536 row: e column: 07
Seq primer: Swfor Forward
Class: methylation filtered
High quality sequence stop: 694.
Location/Qualifiers
1. 694
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATX623"
/db_xref="taxon:4558"
/clone="11137783"
/clone_lib="Sorghum methylation filtered library (LibID:
104)"
/note="Organ: leaf; Vector: pBSK(-); Site: 1: HincII; DNA
prepared from purified nuclei was randomly sheared,
end-repaired, size fractionated to enrich for the 0.5 to 5
kb fraction, ligated into HincII-digested pBSK(-) vector
and electroporated into E. coli cells. This is a
methylation filtered library."
ORIGIN
Query Match 13.4%; Score 40.2; DB 13; Length 694;
Best Local Similarity 52.7%; Pred. No. 10;
Matches 87; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 13 CTAATGCTAGTACGACTACGATGAGACTTGAAATATTTCTTCAATTAACAGCTA 72
Dy 529 CTCAAGCGGAATATTACTATCATATAATTAATAGTACTAGATTGTGATGCTT 470
QY 73 ACTCCCAATGTACCTAACTAGGGGAGTTTATCATGAAGAAATGTTAAATGCTTAA 132
Dy 469 CAACCCGAGTTGACCCCTGTTTGAGAGTTTATTATGCTCTAAAGATCAAGTACTT 410
QY 133 TTCAATAAAGAGTAACGAGTGTGCAACATCTGTGTAATAA 177
Dy 409 TTCAAAATAAAGAAATTCATCTTCGAGTAGTATGTCATTA 365
RESULT 16
CC272514/c
LOCUS
DEFINITION
CC272514 1172 bp DNA linear GSS 13-MAY-2003
CH261-91F1 RM1.1 CH261 Gallus gallus genomic clone CH261-91F1,
genomic survey sequence.
ACCESSION
CC272514

```

```

VERSION
KEYWORDS
SOURCE
ORGANISM
CC272514.1 GI:30627693
GSS.
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 1172)
REFERENCE
AUTHORS
Kremitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
TITLE
JOURNAL
COMMENT
Gallus BAC End Reads
Unpublished (2003)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: RM1 TAGCACTCATATAGGAGGA
Class: BAC ends
High quality sequence start: 115
High quality sequence stop: 316.
Location/Qualifiers
1. 1172
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-91F1"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: http://www.chori.org/bacpac"
ORIGIN
Query Match 13.4%; Score 40.2; DB 12; Length 1172;
Best Local Similarity 54.4%; Pred. No. 11;
Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 90 AACTAGGGGATTTATCATGAAGAAATGTTAAATGCTTAATTTTCAATTAAGAGTAAC 149
Dy 950 AAAAATGGGTGAACAAACAAAGACATGTTTCTTATTAATTTGCAAGTACACTAGAA 891
QY 150 CAGAGTGTGCAACATCTGTGTTAATAACCGCAAACTTCAATCACTATAGCTGTAGTA 209
Dy 890 AAGATTTTGAACATCGGTTTACAAATCGGCCATTTTTTTTCCCAAGGAAAAA 831
QY 210 GAGTGCAATCTGCAAGGATCCAGAGTAA 238
Dy 830 AACGAATTAACAGTAGTAGAAGTAA 802
RESULT 17
CC288289/c
LOCUS
DEFINITION
CD288289 543 bp mRNA linear EST 27-MAY-2003
3.P14.abd POE14 (Day_14_pregnant_ovine_endometrium) Ovis aries
cDNA, mRNA sequence.
ACCESSION
CD288289
VERSION
CD288289.1 GI:31086332
KEYWORDS
EST.
SOURCE
Ovis aries (sheep)
ORGANISM
Ovis aries
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Caprinae; Ovis.
1 (bases 1 to 543)
REFERENCE
AUTHORS
Gray, C.A., Abbey, C.A., Beremand, P.D., Choi, Y., Farmer, J.L.,
Adelson, D.L., Thomas, T.L., Bazer, F.W. and Spencer, T.B.
TITLE
Identification of Endometrial Genes Regulated by Early Pregnancy,
Progesterone, and Interferon Tau in the Ovine Uterus
JOURNAL
Biol. Reprod. 74 (2), 383-394 (2006)
PUBMED
16251498

```



```

CJ381709      745 bp  mRNA  linear  EST 15-JUN-2005
LOCUS         CJ381709 Molgula tectiformis unpublished cDNA library, mixture of
DEFINITION    gastrula and neurula Molgula tectiformis cDNA clone mtga016g03 3',
               mRNA sequence.
ACCESSION     CJ381709
VERSION       CJ381709.1 GI:67822106
KEYWORDS      EST.
SOURCE        Molgula tectiformis
ORGANISM      Molgula tectiformis
               Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
REFERENCE     Stolidobranchia; Molgulidae; Molgula.
AUTHORS       Gyoja,F., Satou,Y. and Satoh,N.
TITLE         Expressed genes in Molgula tectiformis
JOURNAL       Unpublished (2005)
COMMENT       Contact: Tadasu Shin-1
               Center For Genetic Resource Information
               National Institute of Genetics
               1111 Yata, Mishima, Shizuoka 411-8540, Japan
               Tel: 81-559-81-6856
               Fax: 81-559-81-6855
               Email: tshini@genes.nig.ac.jp
               When you want to obtain this EST clone, please send an e-mail to
               Nori Satoh (satoh@ascidian.zool.kyoto-u.ac.jp) and its cc to Yutaka
               Satou. (yutaka@ascidian.zool.kyoto-u.ac.jp).
               Location/Qualifiers
                 1. 745
                   :       /organism="Molgula tectiformis"
                   :       /mol_type="mRNA"
                   :       /db_xref="taxon:30286"
                   :       /clone="mtga016g03"
                   :       /tissue_type="whole animal"
                   :       /dev_stage="mixture of gastrula and neurula"
                   :       /clone_lib="Molgula tectiformis unpublished cDNA library,"
                   :       : mixture of gastrula and neurula"

ORIGIN
Query Match      13.2%; Score 39.6; DB 5; Length 745;
Best Local Similarity 53.2%; Pred. No.15;
Matches 84; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY      4  ATTGACAGGCTAAATGCTTAAGTCACTACGATGAGACTTGAATATATCTTCATTGAAAT 63
DB      80  AATGAAATAATGAATGCTCAGTTTATGGAAGCAACATTTATTTAACTAACTTAAAAAC 139

QY      64  AATCAGCTACTCCCAATTTGACCTTAACCTAGGGAGTTTATCATGAGAAATGTTTAA 123
DB      140  AAATCTGTAACCTAATAATTTTGACATTTGTTGGTTTATGTATGGTAAAGTAAAC 199

QY      124  TGGTTAATTTTCAAATAAGAAAGTAACAGAGTGTGCA 161
DB      200  TGGTTAAGTTTCAGAAACAGTTAATGGCAGTAGTGA 237

RESULT 21
CJ385948      810 bp  mRNA  linear  EST 15-JUN-2005
LOCUS         CJ385948 Molgula tectiformis unpublished cDNA library, mixture of
DEFINITION    gastrula and neurula Molgula tectiformis cDNA clone mtga030j02 3',
               mRNA sequence.
ACCESSION     CJ385948
VERSION       CJ385948.1 GI:67826345
KEYWORDS      EST.
SOURCE        Molgula tectiformis
ORGANISM      Molgula tectiformis
               Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
REFERENCE     Stolidobranchia; Molgulidae; Molgula.
AUTHORS       Gyoja,F., Satou,Y. and Satoh,N.
TITLE         Expressed genes in Molgula tectiformis
JOURNAL       Unpublished (2005)
COMMENT       Contact: Tadasu Shin-1

```

Consortium (IBMC) by Embrapa Recursos Geneticos e Biotecnologia with financing from Conselho Nacional de Desenvolvimento Cientifico e Tecnologico (CNPq), Brazil.
Plate: 507 row: D column: 02

Class: BAC ends
High quality sequence stop: 539.

```
Location/Qualifiers
1. .539
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_507D02"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3"
Hereford bull L1 Domino b
library (Male) produced b
```

ORIGIN

Query Match 13.1%; Score 39.4; DB 13; Length 539;
Best Local Similarity 57.9%; Pred. No. 16;
Matches 70; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

37 AGACTTGAAATATTCTTCAATTTGAAATAAACAGCTAACTCCCAAATTGTACCTAAGG 96

— — — — —

97 GGAGTTTATCATGAGAAATGTTTAAATGCTTAATTTCAAATAAGAAAGTAACCAGAGTG 156

91 C G A C T C C C T C A T C A A C T A A T T T T T A A A T C A T C A T A A C C A C A T T A T A T A T A G C T C A T G C 150

Qy 157 T 157

151 T 151

| RESULT | 23 | 758 bp | DNA | linear | GSS 03-JAN-2006 |
|------------|---------------------------------------------------|--------|-----|--------|-----------------|
| DU930095 | | | | | |
| LOCUS | | | | | |
| DEFINITION | | | | | |
| | DU930095 | | | | |
| | 349135 Tomato ECoRI BAC Library | | | | |
| | clone SL ECoRI0127D09 5. genomic survey sequence. | | | | |

```

FEATURES
source
1. .758
Location/Qualifiers
high quality sequence stop. 345.
/organism="Lycomersicon esculentum"

```

```
/mol_type="genomic DNA"  
/cuiovar="Heinz 1706"  
/db_xref="taxon:4981"  
/clone="SL_ECORI012709"  
/lab_host="E. coli"  
/clone_lib="Tomato ECORI BAC Library"  
/note="#vector: unk; Site 1: ECORI"
```

| | | | | |
|-----------------------|--------|---------------|--------|----------------|
| Query Match | 13.1%; | Score 39.4; | DB 14; | Length 758; |
| Best Local Similarity | 55.5%; | Pred. No. 17; | | |
| Matches | 76; | Conservative | 0; | Mismatches 61; |
| | | | | Indels 0; |
| | | | | Gaps 0; |

QY 10 AGGCTAAATGCTAAGTGACTACGATGAAGACTTGAAAATATTCTTCATTTGAAATAAACAG 69

438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213

7 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

QY
70 CTAACCTCCCAAAATTGTACCTTAACCTAGGGGAGTCTATCATGAAGAAATGTTAAATGCTTA 123

130 ATTTTCAAATAGAACT 146

5

| | | | | |
|------------|---------------------------------------------------------------------------------|--------|-----|--------|
| RESULT 24 | | | | |
| CG804676/c | | | | |
| LOCUS | CG804676 | 547 bp | DNA | linear |
| DEFINITION | 1118952E06.y1.l118 - RescueMu Grid S Zea mays genomic, genomic survey sequence. | | | |

FEATURES
source

| FEATURES | source |
|------------------------------------|------------------------------------|
| 1. <i>Age</i> | 1. <i>Age</i> |
| 2. <i>Gender</i> | 2. <i>Gender</i> |
| 3. <i>Marital status</i> | 3. <i>Marital status</i> |
| 4. <i>Education</i> | 4. <i>Education</i> |
| 5. <i>Income</i> | 5. <i>Income</i> |
| 6. <i>Occupation</i> | 6. <i>Occupation</i> |
| 7. <i>Religion</i> | 7. <i>Religion</i> |
| 8. <i>Political affiliation</i> | 8. <i>Political affiliation</i> |
| 9. <i>Health status</i> | 9. <i>Health status</i> |
| 10. <i>Travel history</i> | 10. <i>Travel history</i> |
| 11. <i>Family size</i> | 11. <i>Family size</i> |
| 12. <i>Home ownership</i> | 12. <i>Home ownership</i> |
| 13. <i>Employment status</i> | 13. <i>Employment status</i> |
| 14. <i>Language spoken</i> | 14. <i>Language spoken</i> |
| 15. <i>Religious observance</i> | 15. <i>Religious observance</i> |
| 16. <i>Political participation</i> | 16. <i>Political participation</i> |
| 17. <i>Health insurance</i> | 17. <i>Health insurance</i> |
| 18. <i>Travel frequency</i> | 18. <i>Travel frequency</i> |
| 19. <i>Family income</i> | 19. <i>Family income</i> |
| 20. <i>Home location</i> | 20. <i>Home location</i> |
| 21. <i>Employment sector</i> | 21. <i>Employment sector</i> |
| 22. <i>Language proficiency</i> | 22. <i>Language proficiency</i> |
| 23. <i>Religious affiliation</i> | 23. <i>Religious affiliation</i> |
| 24. <i>Political ideology</i> | 24. <i>Political ideology</i> |
| 25. <i>Health care access</i> | 25. <i>Health care access</i> |
| 26. <i>Travel mode</i> | 26. <i>Travel mode</i> |
| 27. <i>Family structure</i> | 27. <i>Family structure</i> |
| 28. <i>Home type</i> | 28. <i>Home type</i> |
| 29. <i>Employment duration</i> | 29. <i>Employment duration</i> |
| 30. <i>Language fluency</i> | 30. <i>Language fluency</i> |
| 31. <i>Religious practice</i> | 31. <i>Religious practice</i> |
| 32. <i>Political engagement</i> | 32. <i>Political engagement</i> |
| 33. <i>Health care utilization</i> | 33. <i>Health care utilization</i> |
| 34. <i>Travel frequency</i> | 34. <i>Travel frequency</i> |
| 35. <i>Family income</i> | 35. <i>Family income</i> |
| 36. <i>Home location</i> | 36. <i>Home location</i> |
| 37. <i>Employment sector</i> | 37. <i>Employment sector</i> |
| 38. <i>Language proficiency</i> | 38. <i>Language proficiency</i> |
| 39. <i>Religious affiliation</i> | 39. <i>Religious affiliation</i> |
| 40. <i>Political ideology</i> | 40. <i>Political ideology</i> |
| 41. <i>Health care access</i> | 41. <i>Health care access</i> |
| 42. <i>Travel mode</i> | 42. <i>Travel mode</i> |
| 43. <i>Family structure</i> | 43. <i>Family structure</i> |
| 44. <i>Home type</i> | 44. <i>Home type</i> |
| 45. <i>Employment duration</i> | 45. <i>Employment duration</i> |
| 46. <i>Language fluency</i> | 46. <i>Language fluency</i> |
| 47. <i>Religious practice</i> | 47. <i>Religious practice</i> |
| 48. <i>Political engagement</i> | 48. <i>Political engagement</i> |
| 49. <i>Health care utilization</i> | 49. <i>Health care utilization</i> |
| 50. <i>Travel frequency</i> | 50. <i>Travel frequency</i> |
| 51. <i>Family income</i> | 51. <i>Family income</i> |
| 52. <i>Home location</i> | 52. <i>Home location</i> |
| 53. <i>Employment sector</i> | 53. <i>Employment sector</i> |
| 54. <i>Language proficiency</i> | 54. <i>Language proficiency</i> |
| 55. <i>Religious affiliation</i> | 55. <i>Religious affiliation</i> |
| 56. <i>Political ideology</i> | 56. <i>Political ideology</i> |
| 57. <i>Health care access</i> | 57. <i>Health care access</i> |
| 58. <i>Travel mode</i> | 58. <i>Travel mode</i> |
| 59. <i>Family structure</i> | 59. <i>Family structure</i> |
| 60. <i>Home type</i> | 60. <i>Home type</i> |
| 61. <i>Employment duration</i> | 61. <i>Employment duration</i> |
| 62. <i>Language fluency</i> | 62. <i>Language fluency</i> |
| 63. <i>Religious practice</i> | 63. <i>Religious practice</i> |
| 64. <i>Political engagement</i> | 64. <i>Political engagement</i> |
| 65. <i>Health care utilization</i> | 65. <i>Health care utilization</i> |
| 66. <i>Travel frequency</i> | 66. <i>Travel frequency</i> |
| 67. <i>Family income</i> | 67. <i>Family income</i> |
| 68. <i>Home location</i> | 68. <i>Home location</i> |
| 69. <i>Employment sector</i> | 69. <i>Employment sector</i> |
| 70. <i>Language proficiency</i> | 70. <i>Language proficiency</i> |
| 71. <i>Religious affiliation</i> | 71. <i>Religious affiliation</i> |
| 72. <i>Political ideology</i> | 72. <i>Political ideology</i> |
| 73. <i>Health care access</i> | 73. <i>Health care access</i> |
| 74. <i>Travel mode</i> | 74. <i>Travel mode</i> |
| 75. <i>Family structure</i> | 75. <i>Family structure</i> |
| 76. <i>Home type</i> | 76. <i>Home type</i> |
| 77. <i>Employment duration</i> | 77. <i>Employment duration</i> |
| 78. <i>Language fluency</i> | 78. <i>Language fluency</i> |
| 79. <i>Religious practice</i> | 79. <i>Religious practice</i> |
| 80. <i>Political engagement</i> | 80. <i>Political engagement</i> |
| 81. <i>Health care utilization</i> | 81. <i>Health care utilization</i> |
| 82. <i>Travel frequency</i> | 82. <i>Travel frequency</i> |
| 83. <i>Family income</i> | 83. <i>Family income</i> |
| 84. <i>Home location</i> | 84. <i>Home location</i> |
| 85. <i>Employment sector</i> | 85. <i>Employment sector</i> |
| 86. <i>Language proficiency</i> | 86. <i>Language proficiency</i> |
| 87. <i>Religious affiliation</i> | 87. <i>Religious affiliation</i> |
| 88. <i>Political ideology</i> | 88. <i>Political ideology</i> |
| 89. <i>Health care access</i> | 89. <i>Health care access</i> |
| 90. <i>Travel mode</i> | 90. <i>Travel mode</i> |
| 91. <i>Family structure</i> | 91. <i>Family structure</i> |
| 92. <i>Home type</i> | 92. <i>Home type</i> |
| 93. <i>Employment duration</i> | 93. <i>Employment duration</i> |
| 94. <i>Language fluency</i> | 94. <i>Language fluency</i> |
| 95. <i>Religious practice</i> | 95. <i>Religious practice</i> |
| 96. <i>Political engagement</i> | 96. <i>Political engagement</i> |
| 97. <i>Health care utilization</i> | 97. <i>Health care utilization</i> |
| 98. <i>Travel frequency</i> | 98. <i>Travel frequency</i> |
| 99. <i>Family income</i> | 99. <i>Family income</i> |
| 100. <i>Home location</i> | 100. <i>Home location</i> |
| 101. <i>Employment sector</i> | 101. <i>Employment sector</i> |
| 102. <i>Language proficiency</i> | 102. <i>Language proficiency</i> |
| 103. <i>Religious affiliation</i> | |

ampicillin."

| ORIGIN | Query Match | Score | DB | Length |
|--------|--------------------------|---------------|------------|---------|
| | Best Local Similarity | 12.9%; | 38.8; | 547; |
| | Pred. No. 23; | 50.5%; | | |
| | Matches 94; Conservative | 0; Mismatches | 92; Indels | 0; Gaps |

Q7 4 ATTGACAGCCTAAATGCTAAGTGAAGTCTACGATGAAGACTTGGAAATATCTTCATTTGAAAT 63

DB 363 ATACATATGTATAATTTTATTTATCTCAATATATATAACACCTCGAGCATGATGTTTATACTCGTT 304

Q7 64 A A A C A G C T A A C T C C C A A A T T G T A C C T A A C T A G G G A G T T T A T C A T G A A A A T G T T T A A 123

Db 303 A G A C A C T A C G C G C G T T T T T T A C A T T T T A C A T T T T A C A T T T T A C A T T T T A A G A A A T G T A A G G A 244

Q1 124 TGCCTAATTTTCAATAAGAGTAAACAGAGTGTTCGAAATGCTGTTAAATAACCGAC 183

DB 243 TTTTFTTTTGTAAAGAGATGATGTAGGACGCCCTGTGAACTAGTGTGTTTAAGTACAGTAG 184

| | | | |
|----|-----|-------|-----|
| Q7 | 184 | AACTT | 189 |
| | | | |
| | | | |
| D8 | 183 | AGATT | 178 |

RESULT 25
BH929322/c

DEFINITION
cd102d08.b1 B.oleracea002 Brassica oleracea genomic, linear
sequence.
ACCESSION
BH929322

KEYWORDS
SOURCE
ORGANISM

1 (bases 1 to 690)

REFERENCE

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosoids II; Brassicales; Brassicaceae; Brassica.

TITLE Whole genome shotgun reads from *Brassica oleracea*
UNPUBLISHED (2002)

Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu

Seq primer: -21uppt forward
Class: shotgun
High quality sequence start: 55

| FEATURES | source |
|---------------------|-------------------------------|
| Location/Qualifiers | 1. .890 |
| | /organism="Brassica oleracea" |

```

cc=GenomeData
db_xref="taxon:3712"
clone_lib="B.Oleracea002"
notes="Vector: pOTw13; Whole genome shotgun library from

```

proving that the parasites from a crude nuclear prep using Brassica oleracea 10100DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the

ORIGIN
notary prepared at washington university genomic sequencing center."

Best, Local Similarity 50.5%; Pred. No. 24; Mismatches 92; Indels 0; Gaps 0;
Matches 94; Conservative 0;

[illegible]

| | | | |
|----|-----|-----------------------------------------------------------------|-----|
| Qy | 98 | GAGCTTTATCATGAAGAAATGTTTTAAATGCCTTAATTTTCAAATAAGAAGTAACACAGAGTG | 157 |
| | | | |
| Db | 164 | TAATAAATAGCAATTAATATTTTAGATTTTACAATTTAAAATTTATC AAAATACTGTGATT | 105 |
| | | | |
| Qy | 158 | TGCAACATGCTGTTAAATAAACCCGACAACTTCCAATCACTATAGTCGTGTAGTAGAGTGCAT | 217 |
| | | | |
| Db | 104 | ATTATTTTAAAAATAAAATACAAATAACATATTTTAAATTATTAGATTTTTATAAAGAGATT | 45 |
| | | | |
| Qy | 218 | TC TGCA | 223 |
| | | | |
| Db | 44 | CCACCA | 39 |

| | |
|------------|--------------------------------------------------------------------|
| RESULT | 26 |
| BF336378 | |
| LOCUS | BF336378 359 bp mRNA linear EST 22-NOV-2000 |
| DEFINITION | CW3-CT0510-010900-319-c09 CT0510 Homo sapiens CDNA, mRNA sequence. |
| ACCESSION | BF336378 |
| VERSION | BF336378.1 GI:11307126 |
| KEYWORDS | EST. |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens |

REFERENCE

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P., Bucher, P., Jorgensen, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

1. (bases 1 to 359)

1. (bases 1 to 359)

| TITLE | JOURNAL | PUBLISHED | COMMENT |
|-----------------------------------------------------------------------------------------------------|-------------------------------|--------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Shogren, R. A. Shogren sequencing of the human transcriptome with ORF expressed sequence tags | Proc. Natl. Acad. Sci. U.S.A. | 97 (7), 3491-3496 (2000) | 10737800 Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil |

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?ti=CM3&t2=CM3-CT0510-010900-319-c09&t3=2000-09-01&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 3
 High quality sequence stop: 359.
 Location/Qualifiers

```

FEATURES
source
    location/Qualifiers
    1..359
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /dev_stage="Adult"
    /clone_lib="CT0510"
    /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
    SmaI; A mini-library was made by cloning products derived
    from ORESTES PCR (U.S. Letters Patent application No.
    196,716 - Ludwig Institute for Cancer Research) profiles
    into the pUC 18 vector. Reverse transcription of tissue
    mRNA and cDNA amplification were performed under low
    stringency conditions."

```

| ORIGIN | stringency conditions: |
|-----------------------|---------------------------------------------------------------------|
| Query Match | 12.9%; Score 38.6; DB 7; Length 359; |
| Best Local Similarity | 56.5%; Pred. NO. 25; |
| Matches | 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1; |
| QV | 41 TTGAAATATCTTCATTGTGAAATAACAGCTAACTCCCAAATGTGTACCTAACTAGGGGAG 100 |

| | | | |
|-----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------|--------------------|
| Db | 170 | TTGAAATTATTTTAAAGCTTAAAAATCTGCTGGTTGCAAAATTGTATATACCAATTATATC | 229 |
| Qy | 101 | TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAAGTAACACAGAGTGTG | 159 |
| Db | 230 | TATATTATTAACCTCTAAGTATTTTGGCTGATATGAATATARAAGTAAACACAGGTTA | 289 |
| Qy | 160 | CAACATGCTGTTAAATAACCGCAAACTTCAATCACTATA | 200 |
| Db | 290 | CAGTTGTGCAAAACAAAACAAACATACCTAAATTTACAATA | 330 |
| RESULT 27 | | | |
| BZ049137/c | | | |
| LOCUS | BZ049137 | 759 bp | DNA linear |
| DEFINITION | Jnr21g08_b1 B.oleracea001 Brassica oleracea genomic, genomic survey | | GSS 09-OCT-2002 |
| ACCESSION | BZ049137 | | |
| VERSION | BZ049137.1 | GI:23647330 | |
| KEYWORDS | GSS. | | |
| SOURCE | Brassica oleracea | | |
| ORGANISM | Brassica oleracea | | |
| REFERENCE | Ekaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica. | | |
| AUTHORS | 1 (bases 1 to 759) Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T., Nash,W., Rabinowicz,P.D. and Wilson,R.K. | | |
| TITLE | Whole genome shotgun reads from Brassica oleracea | | |
| JOURNAL | Unpublished (2002) | | |
| COMMENT | Contact: Richard K. Wilson Genome Sequencing Center Washington University School of Medicine Email: submissions@wustl.edu Plate: jnr21 row: g column: 08 Seq primer: -21UPPOT forward Class: shotgun High quality sequence start: 24 High quality sequence stop: 551. | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..759 | | |
| | /organism="Brassica oleracea" | | |
| | /mol_type="genomic DNA" | | |
| | /db_xref="taxon:3712" | | |
| | /clone_lib="B.oleracea001" | | |
| | /note="vector: pOTW13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea Tol000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center." | | |
| ORIGIN | | | |
| Query Match | 12.9%; | Score 38.6; | DB 11; Length 759; |
| Best Local Similarity | 48.4%; | Pred. No. 27; | |
| Matches 107; | Conservative 0; | Mismatches 114; | Indels 0; Gaps 0; |
| Qy | 38 | GACTTGGAAATATCTCTCATTTGGAATAAACACAGCTAACTCCCAAATTTGCTACTTAAGG | 97 |
| Db | 394 | GATTTGCTGATATTTTAAAAAATTTTCAAAAATTTAACTATACATTAATTTAAATATA | 335 |
| Qy | 98 | GAGTTTATCATGAAGAAATGTTTAAATGCTTTAAATTTTCAATAAGAAGTAAACAGAGTGT | 157 |
| Db | 334 | TAATAAAATAGAAATAAATATTTAGATTTTACAATTTGAAATTTATCAAAATCTGTGAT | 275 |
| Qy | 158 | TGCAACATGCTGTTAAATAAACCCACACAACTTCAATCACTAGCTGTAGTAGAGTGCAT | 217 |
| Db | 274 | ATTAATTTTAAATAAAATACAAATACATATTTTAATTAATTAATTTTATAAAATATAT | 215 |
| Qy | 218 | TCTCAAGGATCCACAGAGTACACAGTATTTTGGAAATGCAA | 258 |
| Db | 214 | TATAGTAAATATATAAATGATATAAATATATATAATAAATTACAA | 174 |

availability, please contact Pieter de Jong (pdejong@chori.org). Clones may be purchased from BACPAC Resources (<http://BACPACresources.chori.org>). This work was undertaken as part of the International Swine Genome Sequencing Consortium by University of Illinois at Urbana Champaign, USA with funds provided by grant No. AG2002-34480-11828 from USDA-CSREES and AG2001-35205-09965 from USDA/NRI (Livestock Genome Sequencing Initiative).
 Plate: 260 row: K column: 19
 Seq primer: SP6
 Class: BAC ends.

Location/Qualifiers
 1...683
 /organism="Sus scrofa"
 /mol_type="genomic DNA"
 /strain="four pigs (breed: 37.5% Yorks Landrace and 25% Meishan)"
 /db_xref="taxon:9823"
 /clone="RPC144_260K19"
 /sex="male"
 /cell_type="blood"
 /clone_lib="RPC1-44"
 /note="Vector: pTARBAC2; Site 1: EcoRI; Site 2: EcoRI; porcine male BAC library produced by Pieter de Jong"

FEATURES

source

ORIGIN

Query Match 12.8%; Score 38.4; DB 13; Length 683;
 Best Local Similarity 52.5%; Pred. No. 31;
 Matches 84; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

21 TAAGTACTAGTGAAGACTTGAATATCTTCATTTGAAATAAACAGCTAACTCCAA 80
 |||||
 305 TTATGGCAGCATGAGAAGAGGAAATAGGTGTTTGAATAATTCATTATCTGGCAT 246
 |||||
 81 ATTGTACTACTAGGGAGTTTATCATGAGAATGTTTAATGCTTAATTTCAATA 140
 |||||
 245 TTATCCCTCGATTACGTATTTTAGAGCAGACTTTTAAAGAGCTTAAATTTAAAGGA 186
 |||||

141 AGAAGTACCAAGAGTGTGCAACATGCTGTTAAATAACCC 180
 |||||
 185 TGGACAACACTTTCITTTAAAGGATGTTAATAATCCC 146
 |||||

RESULT 30
 DN376396
 LOCUS
 DEFINITION LIB38529_003_C05_T7_1 LIB38529 Canis familiaris cDNA clone
 LIB38529_003_C05, mRNA sequence.

ACCESSION DN376396.1 GI:60557616
 VERSION EST.
 KEYWORDS
 SOURCE Canis familiaris (dog)
 ORGANISM Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1 (bases 1 to 746)
 Staton,N.R.
 TITLE Direct Submission (Staton,N.R.)
 JOURNAL Unpublished (2005)
 COMMENT Contact: Nick Staton
 Tel: 636 247 6855
 Email: nicholas.r.staton@pfizer.com.

Location/Qualifiers
 1..746
 /organism="Canis familiaris"
 /mol_type="mRNA"
 /db_xref="taxon:9615"
 /clone="LIB38529_003_C05"
 /tissue_type="heart"
 /lab_host="DH10B"
 /clone_lib="LIB38529"
 /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI"

FEATURES

source

ORIGIN

Query Match 12.8%; Score 38.4; DB 9; Length 746;
 Best Local Similarity 54.7%; Pred. No. 31;
 Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

6 TGACAGGCTAAATGCTAAGTACGATGACGATGAGACTTGAATAATTCATTCTTGAATAA 65
 |||||
 590 TAAAGTCAAGCTCANAAGTGGTAAAGAAAAAGAAATTTAACTTTTCTCTGTAAAAA 649
 |||||
 66 ACAGTAACTCCCAATTTGTTACTTAAGTACGAGTGTATCATGAGAAATGTTTAAATG 125
 |||||
 650 ATACCTAACTCCAATGACTTTCTTATCTAAGTATGAGAAATAGCAATAATTCATAAA 709
 |||||
 126 CTTAATTTTCAATAAG 142
 |||||
 710 GTACAGAGTAAATATG 726
 |||||

RESULT 31

DN373570/c

LOCUS

DEFINITION JGI CAAP8099 rev NIH_XGC_tropInt1 Xenopus tropicalis cDNA clone

IMAGE:7714547 3', mRNA sequence.

ACCESSION CX973570.1 GI:58787092

VERSION CX973570

KEYWORDS EST.

SOURCE Xenopus tropicalis (western clawed frog)

ORGANISM Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Xenopodidae; Xenopus; Silurana.

1 (bases 1 to 779)

Richardson,P., Lucas,S., Rokhsar,D., Dettter,J.C., Ng,D.C.,

Brokstein,P. and Lindquist,E.A.

DOE Joint Genome Institute Xenopus tropicalis EST project

Other ESTs: JGI CAAP8099.fwd

Contact: Lindquist,E.A., Richardson,P.

DOE Joint Genome Institute

2800 Mitchell Drive, Walnut Creek, CA 94598, USA

Tel: 925 296 5600

Fax: 925 296 5710

Email: cdna@jgi-psf.org

Tissue Procurement: Robert M. Grainger

cDNA Library Preparation: Bruce Blumberg Laboratory, University of

California, Irvine

DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>

Clone Distribution: I.M.A.G.E. Consortium/LLNL:

<http://image.llnl.gov>

Naming Conventions: EST name is generated by the concatenation of

the JGI Clone Id and the direction of sequencing. The suffix '.rev'

indicates a reverse sequencing read of the insert. It does not

necessarily reflect the orientation of the insert.

Poly-A: Based upon the presence of a run of 14 or more T residues

at the beginning of the sequence, this clone was polyadenylated.

The resulting Poly-T sequence has been removed.

Plate: CAAP 0085 row: e column: 9

High quality sequence stop: 779

POLYA=yes.

Location/Qualifiers

1..779

/organism="Xenopus tropicalis"

/mol_type="mRNA"

/strain="N6 (Nigerian 6th generation inbred)"

/db_xref="taxon:8364"

/clone="IMAGE:7714547"

/tissue_type="Intestine"

/dev_stage="Adult"

/lab_host="ElectroMAX DH10B T1 Phage Resistant cells"

/clone_lib="NIH_XGC_tropInt1"

/note="vector: pCS107; Site 1: EcoRI; Site 2: XhoI; The

library was prepared from 5 ug of poly A+ RNA by oligo-dT

restriction digestion to determine average insert size (96 clones) and by sequencing (-4 96-well plates) to confirm library quality (e.g. the presence of short polyA+ tails, genomic DNA contamination (must be <1%), ribosomal RNA clones (must be <1%), etc.) and to provide a sequence database representing the predominant clones in each library. The clones were sequenced at the University of Missouri-Columbia DNA Core Facility. After production of the libraries, equal numbers of recombinants from each library were pooled to produce a single mixed library (mega-library) for more extensive sequencing. Bioinformatics work was performed by GK Springer's bioinformatics group in Computer Science at the University of Missouri-Columbia. Clone Requests: Requests for clones should be made to the Director of the University of Missouri DNA Core Facility at: bovine@net.missouri.edu. Bonaldo MF, Lennon G, Soares MB. Normalization and Subtraction: Two approaches to facilitate gene discovery. Genome Res. 1996; 6:791-806. Jiang H, Bivens NJ, Ries JE, Whitworth KM, Green JA, Forrester LJ, Springer GK, Didion BA, Mathialagan N, Prather RS, Lucy MC (2001) Constructing cDNA libraries with fewer clones that contain long poly(dA) tails. Biotechniques 31:38-42. Soares MB, MF Bonaldo, P Jelene, L Su, L Lawton, A Efstratiadis. 1994. Construction and characterization of a normalized cDNA library. Proc Natl Acad Sci, 91:9228-9232. TAG TISSUE=Day 14 CL from a pregnant animal TAG_SEQ=CAGGTGAGAT"

ORIGIN

Query Match 12.7%; Score 38.2; DB 8; Length 756;
Best Local Similarity 57.6%; Pred. No. 35;
Matches 87; Conservative 0; Mismatches 63; Indels 1; Gaps 1;
Qy 15 AAATGCTAAGTACGATGACACTTGAATATCTTCATTTGAATTAACAGCTAAC 74
Db |||||
Qy 535 AAATGCTAAGTACATTAACATGATGATGATATTTTATCTGACATCAGATGTT 476
Db |||||
Qy 75 TCCCAATTTGACCT-AACTAGGGGAGTTTATCATGAGAAATGTTTAAATGCTTAATTT 133
Db |||||
Qy 475 TAACGTATATTAACTGAAAGAAAGAAATTTAAGCAGATGATATAATATCTTATTTT 416
Db |||||
Qy 134 TCAATTAAGAGTAACACGAGTGTTCAC 164
Db |||||
Qy 415 TATTAAATACTAAATAAACTGTTCTCTGCA 385
Db |||||

RESULT 33

BH508774
LOCUS BH508774
DEFINITION BHQ34TF BOGQ Brassica oleracea genomic clone BOGQ34, genomic survey sequence.
ACCESSION BH508774
VERSION BH508774.1 GI:17716864
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 759)
AUTHORS Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S., Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
TITLE Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis
JOURNAL Genome Res. 15 (4), 487-495 (2005)
PUBMED 15805490
COMMENT Other GSSs: BOGQ34TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208

Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.

FEATURES
source

Location/Qualifiers
1..759
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BOGQ34"
/clone_lib="BOGQ"
/note="vector: PHOS1, Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"

ORIGIN

Query Match 12.7%; Score 38.2; DB 11; Length 759;
Best Local Similarity 48.4%; Pred. No. 35;
Matches 106; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
Qy 13 CTAATGCTAAGTACGATGACACTTGAATATCTTCATTTGAATTAACAGCTA 72
Db |||||
Qy 29 CCACATTTCTACATTTGATTAAATGATGTTAACTCTCTTCATGGTATATGATCGT 88
Db |||||
Qy 73 ACTCCCAATTTGACCTAACCTAGGGAGTTTATCATGAGAAATGTTTAAATGCTTAATTT 132
Db |||||
Qy 89 TATAATTTTTTATAAATTAATTTAGTAAACCTTCATAATCTCCATTAATTCGGTGAAT 148
Db |||||
Qy 133 TTCAATTAAGAGTAACACGAGTGTTCACATCTGTTAAATACCCGACAACTTCAA 192
Db |||||
Qy 149 AACGACTATAAGTCACAAATTTCTTTGAAAACCGAGATACAAAGGCTCCAAACCTCTT 208
Db |||||
Qy 193 TCACATAGCTGTAGTAGAGTGCAATCTTGCAGGATGCC 231
Db |||||
Qy 209 TCAACATCATCATGTTAGTGCATGATGAACATATGCC 247
Db |||||

RESULT 34

CZ764379/c
LOCUS CZ764379
DEFINITION OC_Ba0119E01.f OC_Ba Oryza coarctata genomic clone OC_Ba0119E01 5', genomic survey sequence.
ACCESSION CZ764379
VERSION CZ764379.1 GI:71204230
KEYWORDS GSS.
SOURCE Oryza coarctata (Porteresia coarctata)
ORGANISM Oryza coarctata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 808)
AUTHORS Kim,H., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and Wing,R.
TITLE Oryza Map Alignment Project - Arizona Genomics Institute
JOURNAL Unpublished (2005)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: twing@genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0119 row: E column: 01
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1..808
/organism="Oryza coarctata"
/mol_type="genomic DNA"

— — — — —

```

Db 2809 ACCATCATGTCACAGAGGAGGAGATTCCTGGCAGAAAAATCAAGAA 2855

RESULT 37
A0836871
LOCUS
DEFINITION HS_5463_A1_F01_T7A_RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plates=1039 Col=1 Row=K, genomic survey sequence.
ACCESSION A0836871
VERSION A0836871
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
1 (bases 1 to 559)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., and
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D., and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end web Server:
http://www.htsc.washington.edu
Plate: 1039 row: K column: 1
Seq primer: T7
Class: BAC ends
High quality sequence stop: 559.
FEATURES
Location/Qualifiers
source
1..559
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=1039 Col=1 Row=K"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"

ORIGIN
Query Match 12.7%; Score 38; DB 11; Length 559;
Best Local Similarity 55.2%; Pred. No. 38;
Matches 74; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 16 AATGCTAGTGACTACGATGAAGACTTGAATATTCCTTCATTTGAAATAACAGCTAACT 75
|||||
Db 205 AATGGTGAGGATGTCGCTGATGTTCTATTATTAATGATCTGTATTAGACAGTTAAGA 264
|||||

Qy 76 CCCAAATGTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAATGCTTAATTTTC 135
|||||
Db 265 TACATATGGTGTGATGTGTGTGTGATTTCATTAATAAGTGTGATTTATTATATTTTC 324
|||||

Qy 136 AATAAAGAAGTAAC 149
|||||
Db 325 TACTATTAAATTAC 338
|||||

```

```

RESULT 38
DW602832/c
LOCUS
DEFINITION CGX162-D05.xld-t SHGC-CGX2 Gasterosteus aculeatus cDNA clone
CGX162-D05 3', mRNA sequence.
ACCESSION DW602832
VERSION DW602832.1
KEYWORDS GI:85192841
SOURCE EST.
ORGANISM Gasterosteus aculeatus (three spined stickleback)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
TITLE Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
JOURNAL Gasterosteidae; Gasterosteus.
COMMENT 1 (bases 1 to 1091)
Kingsley,D.M., Peichel,C., Knecht,A., Balabhadra,S., Grimwood,J.,
Dickson,M., Schmutz,J. and Myers,R.M.
Expressed sequence tags from Gasterosteus aculeatus (2004)
Unpublished (2004)
Contact: Grimwood, Jane
Stanford Human Genome Center
Stanford University School of Medicine
975 S California Ave, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@shgc.stanford.edu
Plate: 162
High quality sequence start: 30
High quality sequence stop: 901.
FEATURES
Location/Qualifiers
source
1..1091
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Conner Creek sticklebacks, WA USA"
/db_xref="taxon:69293"
/clone="CGX162-D05"
/sex="mixed male and female"
/tissue_type="eyes"
/dev stage="adult"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="SHGC-CGX2"
/note="Vector: Express 1; Total and poly A+ RNA was
isolated from the indicated stickleback tissue, and a cDNA
library was constructed in the Express 1 plasmid vector by
Open Biosystems. First strand cDNA synthesis was primed
with an 54 bp linker primer containing an oligodT sequence
preceded by a synthetic NotI site (first strand primer:
5'-GACTAGTCTAGATCGGAGCGGCCCTT)25-3'). Following
second strand synthesis, cDNAs were made blunt at the end
corresponding to the original 5 prime end of mRNA, and
cloned directionally into the NotI and EcoRV sites of
Express 1. Note that the EcoRV site is typically destroyed
in the blunt end cloning, leaving a junction of the form
'xxxATC' (where is ATC is the second half of the EcoRV
site, and xxx is derived from the cDNA sequence). A map of
the Express 1 vector is available at:
http://www.openbiosystems.com/cdna library construction fa
q.php#8 The primary library was transcribed and amplified
in DH10B (T1 phage resistant) bacteria."

```

ORIGIN

```

Query Match 12.7%; Score 38; DB 10; Length 1091;
Best Local Similarity 55.5%; Pred. No. 41;
Matches 71; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 19 GCTAAGTGACTACGATGAAGACTTGAATATTCCTTCATTTGAAATAACAGCTAACTCCC 78
|||||
Db 1067 GGTAAAGCGAACCCNTATGGAACCTGGTGTGTTTGTCTATAGTGACATGACCATTTTACTGCA 1008
|||||

Qy 79 AAATTGTACCTACTAGGGGAGTTTATCATGAAGAAATGTTTAATGTTTAAATTTTCAA 138
|||||
Db 1007 GAAGNGAGCATACCCCTCTTGTCTTTTACCATGCCTCAATGCATTTTATGCTTAGTGTCAA 948
|||||

```

```

QY      139 TAAGAAGT 146
Db      947 GAAGATGT 940

RESULT 39
CZ953670      914 bp DNA linear GSS 11-AUG-2005
DEFINITION   265847 Tomato EcoRI BAC Library Lycopersicon esculentum genomic
              clone SL_EcoRI0057L14 5, genomic survey sequence.
ACCESSION    CZ953670
VERSION      CZ953670.1 GI:72299235
KEYWORDS     GSS.
SOURCE       Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM     Lycopersicon esculentum
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
              asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE    1 (bases 1 to 914)
AUTHORS      Mueller,L.A., Buels,R.M., Wang,Y., Tanksley,S.D., Giovannoni,J.J.,
              Van Eck,J. and Stack,S.
TITLE        BAC end sequencing from three Solanum lycopersicon libraries
JOURNAL      Unpublished (2005)
COMMENT      Other_GSSs: 273838
              Contact: Lukas Mueller
              Tanksley Lab, Dept. of Plant Breeding
              Cornell University
              251 Emerson Hall, Ithaca, NY 14853, USA
              Tel: 607-255-6557
              Fax: 607-255-6683
              Email: sgn-feedback@sgn.cornell.edu
              Plate: 57 row: L column: 14
              Seq primer: T7
              Class: BAC ends
              High quality sequence start: 6
              High quality sequence stop: 818.
              Location/Qualifiers
                source          1..914
                  /organism="Lycopersicon esculentum"
                  /mol_type="genomic DNA"
                  /cultivar="Heinz 1706"
                  /db_xref="taxon:4081"
                  /clone="SL_EcoRI0057L14"
                  /lab_host="E. coli"
                  /clone_lib="Tomato EcoRI BAC Library"
                  /note="Vector: unk; Site_1: EcoRI"

FEATURES
  source          1..914
    Query Match   12.6%; Score 37.8; DB 13; Length 914;
    Best Local Similarity 54.7%; Pred. No. 45;
    Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY      10 AGGCTAATGCTAGTACGATGAGACTTGAAATATTTCTTCATTGAAATAAACAG 69
Db      554 ATGGAATAAACTCAATAAAGCGCTCAAACTCAAACTCAATCATCATATGAGATAAAAG 613

QY      70 CTAATCTCCAAATGTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTA 129
Db      614 TGACCTTCCAAAGCTTACCACACTAATGAAATTTAATTCGCGATAGTTACAAAAATA 673

QY      130 ATTTTCAATAAGAAGT 146
Db      674 TTCACCAATCAAAATT 690

RESULT 40
DU931319      950 bp DNA linear GSS 03-JAN-2006
LOCUS         353827 Tomato EcoRI BAC Library Lycopersicon esculentum genomic
DEFINITION   clone SL_EcoRI0062K06 3, genomic survey sequence.
ACCESSION    DU931319
VERSION      DU931319.1 GI:84252534
KEYWORDS     GSS.
SOURCE       Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM     Lycopersicon esculentum
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
              Pecora; Bovidae; Bovinae; Bos.
REFERENCE    1 (bases 1 to 684)
AUTHORS      Smith,T.P.L., Roberts,A.J., Echtenkamp,S.E., Chitko-McKown,C.G.,
              Wray,J.E. and Keele,J.W.
TITLE        A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL      Unpublished (2003)
COMMENT      Contact: Smith TPL
              USDA, ARS, US Meat Animal Research Center
              PO Box 166, Clay Center, NE 68933-0166, USA
              Tel: 402 762 4366

```

```

SOURCE       Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM     Lycopersicon esculentum
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
              asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE    1 (bases 1 to 950)
AUTHORS      Mueller,L.A., Buels,R.M., Wang,Y., Tanksley,S.D., Giovannoni,J.J.,
              Van Eck,J. and Stack,S.
TITLE        BAC end sequencing from three Solanum lycopersicon libraries
JOURNAL      Unpublished (2005)
COMMENT      Other_GSSs: 269087
              Contact: Lukas Mueller
              Tanksley Lab, Dept. of Plant Breeding
              Cornell University
              251 Emerson Hall, Ithaca, NY 14853, USA
              Tel: 607-255-6557
              Fax: 607-255-6683
              Email: sgn-feedback@sgn.cornell.edu
              Plate: 62 row: K column: 6
              Seq primer: SP6
              Class: BAC ends
              High quality sequence start: 54
              High quality sequence stop: 564.
              Location/Qualifiers
                source          1..950
                  /organism="Lycopersicon esculentum"
                  /mol_type="genomic DNA"
                  /cultivar="Heinz 1706"
                  /db_xref="taxon:4081"
                  /clone="SL_EcoRI0062K06"
                  /lab_host="E. coli"
                  /clone_lib="Tomato EcoRI BAC Library"
                  /note="Vector: unk; Site_1: EcoRI"

ORIGIN
  Query Match   12.6%; Score 37.8; DB 14; Length 950;
  Best Local Similarity 54.7%; Pred. No. 46;
  Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY      10 AGGCTAATGCTAAGTACGATGAGACTTGAAATATTTCTTCATTGAAATAAACAG 69
Db      553 ATGGAATAAACTCAATAAAGCGCTCAAACTCAAACTCAATCATCATATGAGATAAAAG 612

QY      70 CTAATCTCCAAATGTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTA 129
Db      613 TGACCTTCCAAAGCTTACCACACTAATGAAATTTAATTTGAGCATAGTTACAAAAAATT 672

QY      130 ATTTTCAATAAGAAGT 146
Db      673 ATTCACCAAACTAAAT 689

RESULT 41
CZ430880      684 bp mRNA linear EST 25-MAR-2003
LOCUS         606805 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.
DEFINITION    CB430880
ACCESSION     CB430880
VERSION       CB430880.1 GI:29207368
KEYWORDS      EST.
SOURCE        Bos taurus (cattle)
ORGANISM      Bos taurus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
              Pecora; Bovidae; Bovinae; Bos.
REFERENCE    1 (bases 1 to 684)
AUTHORS      Smith,T.P.L., Roberts,A.J., Echtenkamp,S.E., Chitko-McKown,C.G.,
              Wray,J.E. and Keele,J.W.
TITLE        A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL      Unpublished (2003)
COMMENT      Contact: Smith TPL
              USDA, ARS, US Meat Animal Research Center
              PO Box 166, Clay Center, NE 68933-0166, USA
              Tel: 402 762 4366

```



```

Query Match      12.5%; Score 37.6; DB 11; Length 837;
Best Local Similarity 49.5%; Pred. No. 51;
Matches 97; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 43 GAAATATTCCTTCATTTGAAATAAAGACCTAACTCCCAAAATGTACCTAACTAGGGGAGTT 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 202 GAAGCATACATATAGAAATATTAATATACATTAACCTTCATCATAATTTTGGTA 143

QY 103 TATCATGAAGAAATGTTTAAATGCTTAAATTTCAAATTAAGAAGTAACAGAGTGTGCAA 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 142 TAAAGGGAAGAAATTAATAATAATAATATACAAATATAAGAAAAAATAATAATAGGTA 83

QY 163 CATGCTGTTAAATAACCCGACAACCTCAATCATATAGCTGTAGTAGAGTGCATCTGC 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 82 AAATACAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 23

QY 223 AAGATCCAGAGTAA 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22 AATGAATAATAAGTAA 7

RESULT 44
BU355289
LOCUS 603474334F1 CSEQCHN70 Gallus gallus cDNA clone CHEST355d10 5', mRNA
DEFINITION BU355289 991 bp mRNA linear EST 28-NOV-2002
ACCESSION BU355289
VERSION BU355289.1 GI:25863290
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
REFERENCE 1 (bases 1 to 991)
AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.I., Fickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE A Comprehensive Collection of Chicken CDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. .991
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST355d10"
/dev_stage="3c"
/lab_host="DH10B"
/clone_lib="CSEQCHN70"
/notes="Organ: hearts; Vector: pBluescript II KS(+);
Site 1: EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

```

```

FEATURES
source
1. .373
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:543210"
/db_xref="taxon:9606"
/clone="IMAGE:130747"
/sex="Female"
/dev_stage="placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares placenta Nb2HP"
/notes="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
AATCGAAGATTCGCGCGCAGGATTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo. "

```

```

Query Match      12.5%; Score 37.6; DB 3; Length 991;
Best Local Similarity 51.8%; Pred. No. 52;
Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 64 AAACGCTAACTCCCAAAATGTACCTAACTAGGGGAGTTTATCATGAAGAAATGTTTAAA 123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 AAACATCGAATCTACTACTATCTCTTAAAGAGAAAAAGAAAAAGAAAGTTTGAAA 135

QY 124 TGCTTAATTTTCAAATTAAGAAGTAACCAAGAGTGTTCACACATCTGTATAAATACCCGAC 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 TGCTGACTTACATTCGAAAAAACAACACTTTTTCACACATTTTCAACAATTTACACAA 195

QY 184 AAACCTCAATCTATAGCTGTAGTAGAGTGCATTCGCAAGGA 227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 AATATTTCACGCAAAATGGGAAGTTGGGTCTGTTTGGACAGA 239

```

```

RESULT 45
R22138/c
LOCUS yh25b10.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
DEFINITION IMAGE:130747 3', mRNA sequence.
ACCESSION R22138
VERSION R22138.1 GI:776919
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 373)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Treviski,E., Waterston,R., Williamson,A., Wohldmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 2451
High quality sequence stops: 271
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 2451 Std Error: 0.00
Seq primer: -21m13
High quality sequence stop: 271.
Location/Qualifiers
1. .373

```

ORIGIN

| Query Match | 12.5% | Score 37.4 | DB 10 | Length 373 |
|-----------------------|----------------|---------------------------------------------------------------|----------|------------|
| Best Local Similarity | 50.4% | Pred. No. 51 | | |
| Matches 121 | Conservative 0 | Mismatches 112 | Indels 7 | Gaps 1 |
| QY | 44 | AAATATTCTTCATTGAAATAAACAAGCTAACTCCCAAATTGTACTAACTAGGGGAGTTT | 103 | |
| DB | 261 | AGATAATTTTAAATTAACATACACATTTTCTCTTAAAGGGTCATATCTATTAGGCCCTT | 202 | |
| QY | 104 | ATCATGAAGAAATGTTTTAAATGCTTAATTTTCAAATAAAGAGTAACCGAGTGTTTGCAC | 163 | |
| DB | 201 | TCCCTGCTGTAATCTTAAATGATTAATTTGTGAATTTTACTAA-----GTTTCATC | 149 | |
| QY | 164 | ATGCTGTTAAATAACCCGACAAACTTCCAATCACTATAGCTGTAGTAGTGCAATTTCTGCA | 223 | |
| DB | 148 | AAGNTCTTAAAGAGAGAGAAATAAAGATTAATTTATATGTCTAGTCCCAACTTGG | 89 | |
| QY | 224 | AGGATCCGAGATTAACGATATTTTGGAAATGCAATGTTGAACCGACATCTAATTATTC | 283 | |
| DB | 88 | AAGGTAACAACAGTAGTAGTAACCTTTTGGGAGAAATCTTAAAACTACTTTTAAAGAAGC | 29 | |

RESULT 46
AQ321454
LOCUS
DEFINITION
514 bp DNA linear GSS 06-MAY-1999
RPC111-100A20.TV RPCI-11 Homo sapiens genomic clone RPCI-11-100A20,
genomic survey sequence.

| | |
|-----------|---------------------------------------------------------------------|
| ACCESSION | Q0321454 |
| VERSION | Q0321454.1 |
| KEYWORDS | GI:4054188 |
| SOURCE | GSS. |
| ORGANISM | Homo sapiens (human) |
| | Homo sapiens |
| | Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; |
| | Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; |
| | Hominidae; Homo. |
| REFERENCE | 1 (bases 1 to 514) |
| AUTHORS | Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., |
| | Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C. |
| TITLE | Use of human BAC End Sequences for Sequence-Ready Map Building |
| JOURNAL | Unpublished (1998) |
| COMMENT | Contact: Shaying Zhao, William Nierman, Mark Adams |

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tcdb/humgen/bac_end_search/bac_end_search.html

Seq primer: T7
Class: BAC ends.

| FEATURES | source |
|-----------|---------------------------------------------------------------------------------------|
| Crabp. L2 | Location/Qualifiers |
| | 1. .514 |
| | /organism="Homo sapiens" |
| | /mol_type="genomic DNA" |
| | /db_xref="GDB:7538035" |
| | /db_xref="taxon:9606" |
| | /clone="RPC1-11-100A20" |
| | /sex="Male" |
| | /cell_type="Lymphocytes" |
| | /clone_lib="RPC1-11" |
| | /notes="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; RPC111 Human Male BAC Library" |

Query Match 12.5%; Score 37.4; DB 11; Length 514;
Best Local Similarity 70.4%; Pred. No. 54;
Matches 50; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

[illegible]

| | | | | |
|------------|-----------------------------------------------------------------------------------------|--------|-----|--------|
| RESULT 47 | | | | |
| AQ627491 | | | | |
| LOCUS | AQ627491 | 563 bp | DNA | linear |
| DEFINITION | CITBI-E1-2653P1.TF CITBI-E1 Homo sapiens genomic clone 2653P1, genomic survey sequence. | | | |

| ACCESSION | VERSION | KEYWORDS | SOURCE | ORGANISM |
|-----------|------------|------------|-------------------|---------------|
| AO627491 | AO627491.1 | GI:5089883 | GSS. | |
| | | | Homo sapiens | (human) |
| | | | Homo sapiens | |
| | | | Eukaryota; | Metazoa; |
| | | | Chordata; | Cranialata; |
| | | | Vertebrata; | Euteleostomi; |
| | | | Mammalia; | Eutheria; |
| | | | Euarchontoglires; | Primates; |
| | | | Catarrhini; | |
| | | | Hominidae; | Homo. |

REFERENCE
1. (bases 1 to 563)
Zhao, S., Adams, M.D., Nierman, W., Malek, J., Shizuya, H., Simon, M. and Venter, C. *The complete genome of *Chlamydia trachomatis* serovar L2*. *Proc. Natl. Acad. Sci. USA* 94: 12942-12947 (1997).

| JOURNAL TITLE | COMMENT |
|----------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Map Building Use of BAC End Sequences from Caltech Libraries for Sequence-Ready Unpublished (1997) | Other_GSSs: CITBI-EL-2653P1.TR Contact: CHAYING Zhao, William Nierman, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel: 301 838 0200 Fax: 301 838 0208 Email: hbe@tigr.org Clones are available from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html . Seq primer: M13-21 |

```

FEATURES
source
  Location/Qualifiers
    1..563
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /clone="2653p1"
      /sex="male"
      /cell_type="sperm"
      /clone_lib="CITBI-E1"
      /notes="Vector: pBelosBAC11; Site_1: EcoRI; Site_2: EcoRI;
        CalTech Human BAC Library D"

```

| | | | | |
|-----------------------|--------|---------------|--------|----------------------------------|
| Query Match | 12.5%; | Score 37.4; | DB 11; | Length 563; |
| Best Local Similarity | 70.4%; | Pred. No. 54; | | |
| Matches | 50; | Conservative | 0; | Mismatches 21; Indels 0; Gaps 0; |

| | | | |
|----|-----|--------------------------------------------------------------|-----|
| QY | 21 | TAAGTGACTACGATGAAGACTTGAATATCTTCATCTTTGAAATAAACAGCTAACTCCCAA | 80 |
| | | | |
| Db | 141 | TAATTTTCACAAGTAAACATGTAATATCTTGGTTTAAATGAAAAACAGCTAAATCCCAA | 200 |
| | | | |
| QY | 81 | ATTGTACCTAA | 91 |
| | | | |
| Db | 201 | GTTATTGGTAA | 211 |
| | | | |

| | | | | | |
|------------|------------------------------------------------------------------------------------------------|--------|-----|--------|-----------------|
| RESULT 48 | | | | | |
| CW685077 | | | | | |
| LOCUS | CW685077 | 704 bp | DNA | linear | GSS 01-NOV-2004 |
| DEFINITION | OG_BB0043A15.f OG_Bba Oryza glaberrima genomic clone OG_BB0043A15 5', genomic survey sequence. | | | | |

```

ACCESSION CW685077
VERSION CW685077.1 GI:55153667
KEYWORDS GSS.
SOURCE Oryza glaberrima (African rice)
ORGANISM Oryza glaberrima
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
clade; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1. (bases 1 to 704)
AUTHORS Kim,H., Yu,Y., Wisotski,M., Byrne,M., Stum,D., Smart,D., Rao,K.,
Luo,M., Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C.
and Wing,R.
TITLE OMAP
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0043 row: A column: 15
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.
FEATURES
source
1..704
Location/Qualifiers
/organism="Oryza glaberrima"
/mol_type="genomic DNA"
/db_xref="taxon:4538"
/clone="OG BBA0043A15"
/tissue_type="young leaves"
/lab_host="DH10B T1 phage resistant"
/clone_lib="OG BBA"
/note="Vector: pGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match 12.5%; Score 37.4; DB 13; Length 704;
Best Local Similarity 53.8%; Pred. No. 56;
Matches 77; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 63 TAAACAGCTAACTCCCAATTTGACCTAACCTAGGGAGTTTATCATGAGAAATGTTTAA 122
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
351 TATAGAAAAAATACTAATATTTTTCGAATCAAGATAAATTTATATATAAGAAATATTTAA 410
QY 123 ATGCTTAATTTTCAATAAGAGTAAACAGAGTGTTCACATGCTGTTAAATAACCGA 182
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
411 CTATTAAATTTAAAAAAAATTTGATAATATAAATTAATTAATTTTATATAAATAGT 470
QY 183 CAAACTTCAATCACTATAGCTGT 205
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
471 CAACTAGACGAGTTTAACTTT 493

RESULT 49
BZ047105/c 723 bp DNA linear GSS 09-OCT-2002
LOCUS lxf88e06.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
DEFINITION sequence.
ACCESSION BZ047105
VERSION BZ047105.1 GI:23643266
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1. (bases 1 to 723)
AUTHORS Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
Nash,W., Rabinowicz,P.D. and Wilson,R.K.
TITLE Whole genome shotgun reads from Brassica oleracea

JOURNAL Unpublished (2002)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: lbf88@wustl.edu
Plate: lbf88 row: e column: 06
Seq primer: -2lUPpOT forward
Class: shotgun
High quality sequence start: 22
High quality sequence stop: 551.
Location/Qualifiers
1..723
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/note="Vector: pOTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea TO1000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

ORIGIN
Query Match 12.5%; Score 37.4; DB 11; Length 723;
Best Local Similarity 49.5%; Pred. No. 56;
Matches 95; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 20 CTAAGTGACTACGATCAAGACTTTGAATATTTCTTCATTTGAAATAACAGCTAACTCCA 79
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 CAAAGTAATAAATGATTTGAATGTTATATTTATGTAATATATGTAATATATCCCATTA 241
QY 80 AATTGTACCTAAGTAGGGAGTTTATCATGAGAAATGTTTAAATGCTTAATTTCAAT 139
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
240 GAATATACACTCCCATGCTATTTTCTCTAAATATAGATATTAAATAGTAGTATTAGTAT 181
QY 140 AGAAGTACACAGAGTGTTCACATGCTGTTAAATAACCCGACAACTTCAATCACTAT 199
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
180 AAAAGCAAAAAAATAATTTTCAAANTGACCTTTTGACTTTTCAAACCTTAAATAATTAT 121
QY 200 AGCTGTAGTAGA 211
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120 AACTAAATAAA 109

RESULT 50
BZ479314/c 797 bp DNA linear GSS 13-DEC-2002
LOCUS BONHM39TF BO.1.6_2_KB_tot Brassica oleracea genomic clone BONHM39,
DEFINITION genomic survey sequence.
ACCESSION BZ479314
VERSION BZ479314.1 GI:26781712
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1. (bases 1 to 797)
AUTHORS Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
TITLE Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
JOURNAL Genome Res. 15 (4), 487-495 (2005)
PUBMED 15805490
COMMENT Other_GSSs: BONHM39TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org

```

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF

Class: sheared ends.

FEATURES

source

Class. Entered Encls.

1. 797

```

1. .131
/organism="Brassica oleracea"

```

```
/mol type="genomic DNA"
```

```
/mot_type="УПОМОЩЕ"  
/straid="ТО1000РНЗ"
```

```
/BLCLALL="LOI000DH3"  
/db xref="taxon:3712"
```

```

/clone="BONHM39"
/aw_xref="axon:

```

```
/clone="BONHM39"
/clone lib="BO 1.6 2 KB tot"
```

```

/clone_id="BO_1.6_2_KB_tot"
/notes="Vector: pHOSt; Site 1: BstXI: 1.6-2 kb sheared

```

total DNA inserted into pHOS1 using BstXI linkers"

ORIGIN

| Query Match | 12.5% | Score 37.4; | DB 11; | Length 797; |
|-----------------------|-----------------|-----------------------------------------------------------------|-----------|-------------|
| Best Local Similarity | 55.9%; | Pred. No. 57; | | |
| Matches 71; | Conservative 0; | Mismatches 56; | Indels 0; | Gaps 0; |
| QY | 38 | GAC TTGAAATATCTTCATTTGAAATAAACAGCTAACTCCCAAAATGCTACCTAACTAGGG | 97 | |
| | | | | |
| DB | 746 | GAC TTGAAAGAGTCACTATATATAATAGACATCTGATGCCCTTAATTAGGGCTAAGCACTG | 687 | |
| | | | | |
| QY | 98 | GAG TTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAAATAAGAAAGTAACCAAGAGTGT | 157 | |
| | | | | |
| DB | 686 | AAATTTGCTCTAAACCAATTTTCAGAAACACATTTTGTTATGTTGGGGAGTGAACACACAA | 627 | |
| | | | | |
| QY | 158 | TGCACA | 164 | |
| | | | | |
| DB | 626 | TGCAGCA | 620 | |

Search completed: July 17, 2006, 23:22:58
Job time : 2907 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OW nucleic - nucleic search, using sw model

Run on: July 17, 2006, 22:15:53 ; Search time 133 Seconds
(without alignments)
4220.546 Million cell updates/sec

Title: SEQ1-96535C
Perfect score: 300
Sequence: 1 ggcattacaggctaaatgc.....atctgctaggtagaagtat 300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

- 1: /EMC_Celerra_SID83/ptodata/2/ina/1 COMB.seq.*
- 2: /EMC_Celerra_SID83/ptodata/2/ina/5 COMB.seq.*
- 3: /EMC_Celerra_SID83/ptodata/2/ina/6A COMB.seq.*
- 4: /EMC_Celerra_SID83/ptodata/2/ina/6B COMB.seq.*
- 5: /EMC_Celerra_SID83/ptodata/2/ina/7 COMB.seq.*
- 6: /EMC_Celerra_SID83/ptodata/2/ina/H COMB.seq.*
- 7: /EMC_Celerra_SID83/ptodata/2/ina/PTUS COMB.seq.*
- 8: /EMC_Celerra_SID83/ptodata/2/ina/RE COMB.seq.*
- 9: /EMC_Celerra_SID83/ptodata/2/ina/PP COMB.seq.*
- 10: /EMC_Celerra_SID83/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | Query Match | Length | ID | Description |
|--------|-----|-------|-------------|--------|----|---------------------|
| C | 1 | 39.4 | 13.1 | 49164 | 3 | US-09-949-016-12985 |
| C | 2 | 39.4 | 13.1 | 49164 | 3 | US-09-949-016-12986 |
| C | 3 | 39.4 | 13.1 | 49164 | 3 | US-09-949-016-13262 |
| C | 4 | 39.4 | 13.1 | 49164 | 3 | US-09-949-016-13263 |
| C | 5 | 38.4 | 12.8 | 192302 | 3 | US-09-949-016-13270 |
| C | 6 | 36.4 | 12.1 | 832 | 3 | US-09-621-976-2813 |
| C | 7 | 35.6 | 11.9 | 260247 | 3 | US-09-949-016-13358 |
| C | 8 | 35.4 | 11.8 | 48682 | 3 | US-09-949-016-17236 |
| C | 9 | 35.4 | 11.8 | 48682 | 3 | US-09-949-016-17237 |
| C | 10 | 35.4 | 11.8 | 48682 | 3 | US-09-949-016-17238 |
| C | 11 | 35.4 | 11.8 | 48682 | 3 | US-09-949-016-17239 |
| C | 12 | 35.4 | 11.8 | 51022 | 3 | US-09-949-016-17135 |
| C | 13 | 35.4 | 11.8 | 51022 | 3 | US-09-949-016-17136 |
| C | 14 | 35.4 | 11.8 | 51022 | 3 | US-09-949-016-17137 |
| C | 15 | 35.4 | 11.8 | 51022 | 3 | US-09-949-016-17138 |
| C | 16 | 35.4 | 11.8 | 65415 | 3 | US-09-949-016-16699 |
| C | 17 | 35.4 | 11.8 | 86213 | 3 | US-09-949-016-17240 |
| C | 18 | 35.4 | 11.8 | 86213 | 3 | US-09-949-016-17241 |
| C | 19 | 35.4 | 11.8 | 86213 | 3 | US-09-949-016-17242 |
| C | 20 | 35.4 | 11.8 | 86213 | 3 | US-09-949-016-17243 |
| C | 21 | 34.6 | 11.5 | 821 | 3 | US-08-956-171E-156 |
| C | 22 | 34.6 | 11.5 | 821 | 3 | US-08-781-986A-156 |
| C | 23 | 34.2 | 11.4 | 38702 | 3 | US-09-949-016-13788 |

| | | | | | |
|--------------------|----------------------|---|------|---------|---|
| Sequence 16406, A | US-09-949-016-16406 | 3 | 11.3 | 68778 | 3 |
| Sequence 15116, A | US-09-949-016-15116 | 3 | 11.3 | 83462 | 3 |
| Sequence 15638, A | US-09-949-016-15638 | 3 | 11.3 | 69764 | 3 |
| Sequence 714, App | US-09-543-681A-714 | 3 | 11.2 | 2412 | 3 |
| Sequence 15744, A | US-09-949-016-15744 | 3 | 11.1 | 64049 | 3 |
| Sequence 16600, A | US-09-949-016-16600 | 3 | 11.1 | 298336 | 3 |
| Sequence 53968, A | US-09-949-016-53968 | 3 | 11.1 | 601 | 3 |
| Sequence 12025, A | US-09-949-016-12025 | 3 | 11.1 | 89268 | 3 |
| Sequence 13302, A | US-09-949-016-13302 | 3 | 11.1 | 90544 | 3 |
| Sequence 2813, Ap | US-09-621-976-2813 | 3 | 11.0 | 832 | 3 |
| Sequence 3, Appli | US-08-929-329-3 | 3 | 11.0 | 1194 | 3 |
| Sequence 24, Appli | US-09-662-254B-24 | 3 | 11.0 | 50000 | 3 |
| Sequence 13185, A | US-09-949-016-13185 | 3 | 11.0 | 251769 | 3 |
| Sequence 13186, A | US-09-949-016-13186 | 3 | 11.0 | 251769 | 3 |
| Sequence 13187, A | US-09-949-016-13187 | 3 | 11.0 | 266748 | 3 |
| Sequence 13188, A | US-09-949-016-13188 | 3 | 11.0 | 266748 | 3 |
| Sequence 12131, A | US-09-949-016-12131 | 3 | 10.9 | 601 | 3 |
| Sequence 16290, A | US-09-949-016-16290 | 3 | 10.9 | 10783 | 3 |
| Sequence 13659, A | US-09-949-016-13659 | 3 | 10.9 | 23105 | 3 |
| Sequence 12582, A | US-09-949-016-12582 | 3 | 10.9 | 23108 | 3 |
| Sequence 13176, A | US-09-949-016-13176 | 3 | 10.9 | 36577 | 3 |
| Sequence 40, Appli | US-09-596-002-40 | 3 | 10.9 | 119211 | 3 |
| Sequence 14382, A | US-09-949-016-14382 | 3 | 10.9 | 192956 | 3 |
| Sequence 923, App | US-10-002-623-923 | 3 | 10.9 | 456 | 3 |
| Sequence 926, App | US-10-002-623-926 | 3 | 10.9 | 456 | 3 |
| Sequence 64, Appli | US-09-809-545A-64 | 3 | 10.9 | 819 | 3 |
| Sequence 17030, A | US-09-949-016-17030 | 3 | 10.9 | 139049 | 3 |
| Sequence 687, App | US-09-248-796A-687 | 3 | 10.8 | 813 | 3 |
| Sequence 12723, A | US-09-949-016-12723 | 3 | 10.8 | 46823 | 3 |
| Sequence 16352, A | US-09-949-016-16352 | 3 | 10.8 | 46940 | 3 |
| Sequence 3, Appli | US-09-617-594A-3 | 3 | 10.7 | 2007 | 3 |
| Sequence 179050, A | US-10-209-507-3 | 3 | 10.7 | 2007 | 3 |
| Sequence 179051, A | US-09-949-016-179051 | 3 | 10.7 | 601 | 3 |
| Sequence 196473, A | US-09-949-016-196473 | 3 | 10.7 | 601 | 3 |
| Sequence 305, App | US-09-487-558B-305 | 3 | 10.7 | 4071 | 3 |
| Sequence 16908, A | US-09-949-016-16908 | 3 | 10.7 | 89670 | 3 |
| Sequence 2, Appli | US-09-762-311-2 | 3 | 10.7 | 99960 | 3 |
| Sequence 16151, A | US-09-949-016-16151 | 3 | 10.7 | 146401 | 3 |
| Sequence 12896, A | US-09-949-016-12896 | 3 | 10.7 | 451924 | 3 |
| Sequence 17305, A | US-09-949-016-17305 | 3 | 10.7 | 451925 | 3 |
| Sequence 1, Appli | US-08-916-421B-1 | 3 | 10.7 | 1664976 | 3 |
| Sequence 23091, A | US-09-513-999C-23091 | 3 | 10.6 | 411 | 3 |
| Sequence 1651, Ap | US-09-710-279-1651 | 3 | 10.6 | 915 | 3 |
| Sequence 1237, Ap | US-09-134-001C-1237 | 3 | 10.6 | 1989 | 3 |
| Sequence 4273, Ap | US-09-710-279-4273 | 3 | 10.6 | 4155 | 3 |
| Sequence 13703, A | US-09-949-016-13703 | 3 | 10.6 | 67755 | 3 |
| Sequence 16179, A | US-09-949-016-16179 | 3 | 10.6 | 157644 | 3 |
| Sequence 16180, A | US-09-949-016-16180 | 3 | 10.6 | 157644 | 3 |
| Sequence 16297, A | US-09-949-016-16297 | 3 | 10.6 | 421118 | 3 |
| Sequence 3532, A | US-09-949-016-3532 | 3 | 10.5 | 601 | 3 |
| Sequence 62973, A | US-09-949-016-62973 | 3 | 10.5 | 601 | 3 |
| Sequence 903, App | US-09-543-681A-903 | 3 | 10.5 | 1092 | 3 |
| Sequence 910, App | US-09-949-016-910 | 3 | 10.5 | 2526 | 3 |
| Sequence 1818, Ap | US-09-949-016-1818 | 3 | 10.5 | 2530 | 3 |
| Sequence 3, Appli | US-08-213-419B-3 | 3 | 10.5 | 6124 | 3 |
| Sequence 12652, A | US-09-949-016-12652 | 3 | 10.5 | 11858 | 3 |
| Sequence 13560, A | US-09-949-016-13560 | 3 | 10.5 | 11858 | 3 |
| Sequence 14320, A | US-09-949-016-14320 | 3 | 10.5 | 24056 | 3 |
| Sequence 16445, A | US-09-949-016-16445 | 3 | 10.5 | 42094 | 3 |
| Sequence 202, App | US-09-531-120-202 | 4 | 10.5 | 129021 | 4 |
| Sequence 14720, A | US-09-949-016-14720 | 3 | 10.5 | 390890 | 3 |
| Sequence 196410, A | US-09-949-016-196410 | 3 | 10.5 | 601 | 3 |
| Sequence 196412, A | US-09-949-016-196412 | 3 | 10.5 | 601 | 3 |
| Sequence 5116, Ap | US-09-248-796A-5116 | 3 | 10.5 | 1086 | 3 |
| Sequence 4723, Ap | US-09-248-796A-4723 | 3 | 10.5 | 1773 | 3 |
| Sequence 21, Appli | US-09-936-271C-21 | 5 | 10.5 | 8622 | 5 |
| Sequence 16412, A | US-09-949-016-16412 | 3 | 10.5 | 95890 | 3 |
| Sequence 13451, A | US-09-949-016-13451 | 3 | 10.5 | 139562 | 3 |
| Sequence 12, Appli | US-09-276-625-12 | 3 | 10.4 | 208 | 3 |
| Sequence 12, Appli | US-10-054-665A-12 | 3 | 10.4 | 208 | 3 |


```
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13262
LENGTH: 49164
TYPE: DNA
ORGANISM: Human
US-09-949-016-13262

Query Match      13.1%; Score 39.4; DB 3; Length 49164;
Best Local Similarity 55.5%; Pred. No. 0.15;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTGACTACGATGAAGACTTGAATAATCTTCAATTTGAATAAACAGCTAACTCC 77
DB 24145 TACATAGTGAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTAATAATTAA 24086
QY 78 CAAATTGACTACTAGGAGGATTTATCATGAGAGAAATGTTTAAATGCTTAATTTCAA 137
DB 24085 CAGAAGTTATTAAAGTAAAGAGCAACAAAGAGAAATGATGATAAAACAAATAATGTAA 24026
QY 138 ATAAGAGTACCAAGAG 154
DB 24025 CTAATAAATAAGAGAG 24009

RESULT 4
US-09-949-016-13263/c
Sequence 13263, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13263
LENGTH: 49164
TYPE: DNA
ORGANISM: Human
US-09-949-016-13263

Query Match      13.1%; Score 39.4; DB 3; Length 49164;
Best Local Similarity 55.5%; Pred. No. 0.15;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTGACTACGATGAAGACTTGAATAATCTTCAATTTGAATAAACAGCTAACTCC 77
DB 24145 TACATAGTGAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTAATAATTAA 24086
QY 78 CAAATTGACTACTAGGAGGATTTATCATGAGAGAAATGTTTAAATGCTTAATTTCAA 137
DB 24085 CAGAAGTTATTAAAGTAAAGAGCAACAAAGAGAAATGATGATAAAACAAATAATGTAA 24026
QY 138 ATAAGAGTACCAAGAG 154
DB 24025 CTAATAAATAAGAGAG 24009

RESULT 5
US-09-949-016-15270/c
Sequence 15270, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15270
LENGTH: 192302
TYPE: DNA
ORGANISM: Human
NAME/KEY: misc feature
LOCATION: (1) (192302)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15270

Query Match      12.8%; Score 38.4; DB 3; Length 192302;
Best Local Similarity 51.8%; Pred. No. 0.47;
Matches 87; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 74 CTCCTAAATGCTACCTAACTAGGGAGTTTATCATGAGAGAAATGTTTAAATGCTTAATTT 133
DB 66927 CTACTCCATTGGATCTAACTTTGAGATACCAAGAAATAATATTGAGCATATCTCTATAGG 66868
QY 134 TCAATAGAGAGTAACAGAGTGTTCGAACATGCTGTAAATACCCGACAACTTCAAT 193
DB 66867 TTAGATACGGAGTTACATCTGACATATAATTTGAGTGACATAAGAAAGTAGACTGAAAT 66808
QY 194 CACTATAGCTGTAGTAGAGTGCAATTCCTCAAGGATCCAGAGTAACCA 241
DB 66807 ACATACAGCATTTGAGTATAGAGTTTTCAGGGTTCAAGGGTAGCAA 66760

RESULT 6
US-09-621-976-2813
Sequence 2813, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET 054PE2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 2813
LENGTH: 832
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
```



```
RESULT 10
US-09-949-016-17238/c
; Sequence 17238, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17238
; LENGTH: 48682
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17238

Query Match      11.8%; Score 35.4; DB 3; Length 48682;
Best Local Similarity 56.4%; Pred. No. 2.4;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAAGCTTGAATAATCTTTCATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACTATCAGCGAGTAATGCCCATTTGAACATCTTATATAAACAGCTAAAGTATTAACTT 14446

QY 84 GTACCTAACTAGGGAGTTTATCATGAGAAATGTTTAAATGCTTAAATTTCAAATA 140
DB 14445 TATCCTCTTTAAGCTAGAATTACACTTGAAATGTATAATTGATATCTGATCAATTA 14389

RESULT 11
US-09-949-016-17239/c
; Sequence 17239, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17239
; LENGTH: 48682
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17239

Query Match      11.8%; Score 35.4; DB 3; Length 48682;
Best Local Similarity 56.4%; Pred. No. 2.4;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAAGCTTGAATAATCTTTCATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACTATCAGCGAGTAATGCCCATTTGAACATCTTATATAAACAGCTAAAGTATTAACTT 14446

QY 84 GTACCTAACTAGGGAGTTTATCATGAGAAATGTTTAAATGCTTAAATTTCAAATA 140
DB 14445 TATCCTCTTTAAGCTAGAATTACACTTGAAATGTATAATTGATATCTGATCAATTA 14389

RESULT 12
US-09-949-016-17135/c
; Sequence 17135, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17135
; LENGTH: 51022
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17135

Query Match      11.8%; Score 35.4; DB 3; Length 51022;
Best Local Similarity 56.4%; Pred. No. 2.5;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAAGCTTGAATAATCTTTCATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACTATCAGCGAGTAATGCCCATTTGAACATCTTATATAAACAGCTAAAGTATTAACTT 14446

QY 84 GTACCTAACTAGGGAGTTTATCATGAGAAATGTTTAAATGCTTAAATTTCAAATA 140
DB 14445 TATCCTCTTTAAGCTAGAATTACACTTGAAATGTATAATTGATATCTGATCAATTA 14389

RESULT 13
US-09-949-016-17136/c
; Sequence 17136, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17136
; LENGTH: 51022
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17136

Query Match      11.8%; Score 35.4; DB 3; Length 51022;
Best Local Similarity 56.4%; Pred. No. 2.5;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAAGCTTGAATAATCTTTCATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACTATCAGCGAGTAATGCCCATTTGAACATCTTATATAAACAGCTAAAGTATTAACTT 14446

QY 84 GTACCTAACTAGGGAGTTTATCATGAGAAATGTTTAAATGCTTAAATTTCAAATA 140
DB 14445 TATCCTCTTTAAGCTAGAATTACACTTGAAATGTATAATTGATATCTGATCAATTA 14389
```

Db 14505 GTCACTATCAGGCAGTAATGCCATTTTGAACAATCTTATATAAACAAGCTAAGTATTAACCTT 14446
 QY 84 GTACCTAACTAGGCGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAAATA 140
 Db 14445 TATCCTCTTTAAGCTAGAAATACACTTTGAAAATGTATAATGATATCTGATCAATTA 14389

RESULT 14
 US-09-949-016-17137/c
 ; Sequence 17137, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17137
 ; LENGTH: 51022
 ; TYPE: DNA
 ; ORGANISM: Human
 ; ORGANISM: Human
 US-09-949-016-17137

Query Match 11.8%; Score 35.4; DB 3; Length 51022;
 Best Local Similarity 56.4%; Pred. No. 2.5;
 Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
 QY 24 GTGACTACGAGTGAAGCTTGAATATTTCTTCATTTGAATAAACAAGCTAAGTATTAACCTT 83
 Db 14505 GTCACTATCAGGCAGTAATGCCATTTTGAACAATCTTATATAAACAAGCTAAGTATTAACCTT 14446
 QY 84 GTACCTAACTAGGCGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAAATA 140
 Db 14445 TATCCTCTTTAAGCTAGAAATACACTTTGAAAATGTATAATGATATCTGATCAATTA 14389

RESULT 15
 US-09-949-016-17138/c
 ; Sequence 17138, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17138
 ; LENGTH: 51022
 ; TYPE: DNA
 ; ORGANISM: Human
 ; ORGANISM: Human
 US-09-949-016-17138

Query Match 11.8%; Score 35.4; DB 3; Length 51022;
 Best Local Similarity 56.4%; Pred. No. 2.5;
 Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGAGTGAAGCTTGAATATTTCTTCATTTGAATAAACAAGCTAAGTATTAACCTT 83
 Db 14505 GTCACTATCAGGCAGTAATGCCATTTTGAACAATCTTATATAAACAAGCTAAGTATTAACCTT 14446
 QY 84 GTACCTAACTAGGCGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAAATA 140
 Db 14445 TATCCTCTTTAAGCTAGAAATACACTTTGAAAATGTATAATGATATCTGATCAATTA 14389

RESULT 16
 US-09-949-016-16699
 ; Sequence 16699, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 16699
 ; LENGTH: 65415
 ; TYPE: DNA
 ; ORGANISM: Human
 ; ORGANISM: Human
 US-09-949-016-16699

Query Match 11.8%; Score 35.4; DB 3; Length 65415;
 Best Local Similarity 51.6%; Pred. No. 2.7;
 Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
 QY 5 TTGACAGGCTAAATGCTAGTACGATGAAGACTTTGAAATATTTCTTCATTTGAAATA 64
 Db 739 TTCAACAATGATTTCAAAAGTAAAGAACTAGAGAAAAGAAATAAATTTTCATATAAGAACTA 798
 QY 65 AACAGCTAACTCCCAATTTGCTACCTAACTAGGAGGAGTTTATCATGAAGAAATGTTTAAAT 124
 Db 799 AATCACCATTTCACAGAGCCCGACGACTGTGGAACTGCAGAGAGACGACACTGGGAAT 858
 QY 125 GCTTAATTTTCAATAAAGAAAGTAAACAGAGTGTGTGCA 161
 Db 859 GCATTATTACAAATGAATAATTTTAAACAAATCTTTTCA 895

RESULT 17
 US-09-949-016-17240/c
 ; Sequence 17240, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17240
 ; LENGTH: 86213

```
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17240

Query Match      11.8%; Score 35.4; DB 3; Length 86213;
Best Local Similarity 56.4%; Pred. No. 2.9;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAAGACTTGAATAATCTTTCATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACATATCAGCAGTAATGCCATTTGAACATCTTATATAAACAGCTAAAGTATTAACTT 14446

QY 84 GTACTTAAGCTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTCAAATA 140
DB 14445 TATCTCTTTAAGCTAGTAATACACTTGAAATGTATAATGATATCTGATCAATTA 14389

RESULT 18
US-09-949-016-17241/c
; Sequence 17241, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17241
; LENGTH: 86213
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17241

Query Match      11.8%; Score 35.4; DB 3; Length 86213;
Best Local Similarity 56.4%; Pred. No. 2.9;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAAGACTTGAATAATCTTTCATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACATATCAGCAGTAATGCCATTTGAACATCTTATATAAACAGCTAAAGTATTAACTT 14446

QY 84 GTACTTAAGCTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTCAAATA 140
DB 14445 TATCTCTTTAAGCTAGTAATACACTTGAAATGTATAATGATATCTGATCAATTA 14389

RESULT 19
US-09-949-016-17242/c
; Sequence 17242, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17242
; LENGTH: 86213
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17242

Query Match      11.8%; Score 35.4; DB 3; Length 86213;
Best Local Similarity 56.4%; Pred. No. 2.9;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAAGACTTGAATAATCTTTCATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACATATCAGCAGTAATGCCATTTGAACATCTTATATAAACAGCTAAAGTATTAACTT 14446

QY 84 GTACTTAAGCTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTCAAATA 140
DB 14445 TATCTCTTTAAGCTAGTAATACACTTGAAATGTATAATGATATCTGATCAATTA 14389

RESULT 20
US-09-949-016-17243/c
; Sequence 17243, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17243
; LENGTH: 86213
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17243

Query Match      11.8%; Score 35.4; DB 3; Length 86213;
Best Local Similarity 56.4%; Pred. No. 2.9;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAAGACTTGAATAATCTTTCATTTGAAATAAACAGCTAACTCCCAATT 83
DB 14505 GTCACATATCAGCAGTAATGCCATTTGAACATCTTATATAAACAGCTAAAGTATTAACTT 14446

QY 84 GTACTTAAGCTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTCAAATA 140
DB 14445 TATCTCTTTAAGCTAGTAATACACTTGAAATGTATAATGATATCTGATCAATTA 14389

RESULT 21
US-08-956-171E-156/c
; Sequence 156, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
```

| | | | |
|-------------------------------------------------------------------------|--------------------------------------------------------------------|-----------------------|-------------------------------------------------------------------------|
| STREET: 9410 Key West Avenue | APPLICATION NUMBER: 11.5%; Score 34.6; DB 3; Length 821; | Query Match | 11.5%; Score 34.6; DB 3; Length 821; |
| CITY: Rockville | Best Local Similarity 56.6%; Pred. No. 1.2; | Best Local Similarity | 56.6%; Pred. No. 1.2; |
| STATE: Maryland | Matches 64; Conservative 0; Mismatches 49; Indels 0; Gaps 0; | Matches | 64; Conservative 0; Mismatches 49; Indels 0; Gaps 0; |
| COUNTRY: USA | | | |
| ZIP: 20850 | | | |
| COMPUTER READABLE FORM: | | | |
| MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage | | | |
| COMPUTER: HP Vectra 486/33 | | | |
| OPERATING SYSTEM: MSDOS version 6.2 | | | |
| SOFTWARE: ASCII Text | | | |
| CURRENT APPLICATION DATA: | | | |
| APPLICATION NUMBER: US/08/956,171E | | | |
| FILING DATE: 20-Oct-1997 | | | |
| CLASSIFICATION: <Unknown> | | | |
| PRIOR APPLICATION DATA: | | | |
| APPLICATION NUMBER: 60/009,861 | | | |
| FILING DATE: January 5, 1996 | | | |
| APPLICATION NUMBER: 08/781,986 | | | |
| FILING DATE: January 3, 1997 | | | |
| ATTORNEY/AGENT INFORMATION: | | | |
| NAME: Mark J. Hyman | | | |
| REGISTRATION NUMBER: 46,789 | | | |
| REFERENCE/DOCKET NUMBER: PB248P1 | | | |
| TELECOMMUNICATION INFORMATION: | | | |
| TELEPHONE: (240) 314-1224 | | | |
| TELEFAX: (301) 309-8439 | | | |
| INFORMATION FOR SEQ ID NO: 156: | | | |
| SEQUENCE CHARACTERISTICS: | | | |
| LENGTH: 821 base pairs | | | |
| TYPE: nucleic acid | | | |
| STRANDEDNESS: double | | | |
| TOPOLOGY: linear | | | |
| SEQUENCE DESCRIPTION: SEQ ID NO: 156: | | | |
| US-08-956-171E-156 | | | |
| Query Match | 11.5%; Score 34.6; DB 3; Length 821; | Query Match | 11.5%; Score 34.6; DB 3; Length 821; |
| Best Local Similarity | 56.6%; Pred. No. 1.2; | Best Local Similarity | 56.6%; Pred. No. 1.2; |
| Matches | 64; Conservative 0; Mismatches 49; Indels 0; Gaps 0; | Matches | 64; Conservative 0; Mismatches 49; Indels 0; Gaps 0; |
| QY | 105 TCATGAGAAATGTTAAATGCTTAATTTTCAAAATAAGAGTAACACGAGTGTTCACACA 164 | QY | 108 TGAAGAAATGTTAAATGCTTAATTTTCAAAATAAGAGTAACACGAGTGTTCACACATGC 167 |
| DB | 135 TTATTATTATAATTTCCAACTTAGCTTACGATTCAGAAATAAAAAAGTATTAAACA 76 | DB | 35855 TGAAGTGGTAGCGAACTGCTGAAGTTTGAATGAGATCTTTAAACAGCTCTCTCATACCT 35796 |
| QY | 165 TGCTGTTAAATACCCGACAAATTCATCACTATAGCTGTAGTAGAGTGCAT 217 | QY | 168 TGTAAATAAACCCGACAAATTCATCACTATAGCTGTAGTAGAGTGCATTCGCAAGCA 227 |
| DB | 75 CCAGAAATATTGCACACATATTACATCACCATGTTTTTCTCTAATTGAAT 23 | DB | 35795 AACTAAAAAACCTGCCAACCATATTAACTATTCGGGAGATTCTTCCATTCTGGACCTT 35736 |
| RESULT 22 | | QY | 228 TCCAGAGTAACCAAGTATTTTGGAAATGCAATGTTGAAC 266 |
| US-08-781-986A-156/c | | | |
| Sequence 156, Application US/08781986A | | | |
| Patent No. 6737248 | | | |
| GENERAL INFORMATION: | | | |
| APPLICANT: Charles Kunsch | | | |
| TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences | | | |
| NUMBER OF SEQUENCES: 5255 | | | |
| CORRESPONDENCE ADDRESS: | | | |
| ADDRESSEE: Human Genome Sciences, Inc. | | | |
| STREET: 9410 Key West Avenue | | | |
| CITY: Rockville | | | |
| STATE: Maryland | | | |
| COUNTRY: USA | | | |
| ZIP: 20850 | | | |
| COMPUTER READABLE FORM: | | | |
| MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage | | | |
| COMPUTER: HP Vectra 486/33 | | | |
| OPERATING SYSTEM: MSDOS version 6.2 | | | |
| SOFTWARE: ASCII Text | | | |
| CURRENT APPLICATION DATA: | | | |
| APPLICATION NUMBER: US/08/781,986A | | | |
| FILING DATE: | | | |
| CLASSIFICATION: 435 | | | |
| PRIOR APPLICATION DATA: | | | |

Ds 35735 GCATTAGAAAGTTCTGAATGGCTAAGAAAGATGTAC 35697

RESULT 24

US-09-949-016-16406
; Sequence 16406, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16406
; LENGTH: 68778
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(68778)
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-16406

Query Match 11.3%; Score 34; DB 3; Length 68778;
Best Local Similarity 52.9%; Pred. No. 7.1;
Matches 73; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

39 ACTTGAATATCTTCATTTGAAATAACACGCTAACTCCAAATGTACCTAGGAGTTTATCATGAAGAAA 115
38112 ACTTGAATATCTTCATTTGAAATAACACGCTAACTCCAAATGTACCTAGGAGTTTATCATGAAGAAA 4139

99 AGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAAAATAAGAGTAACCAAGAGTGTT 158
38172 TGATCAATAAAGAACAGAGTAAGTTTATAGATATCTGTAATATCTGTAATATCTGTTGTT 38231

159 GCAACATGCTGTTAAATA 176
38232 GGATCAATTAATAAATA 38249

RESULT 25

US-09-949-016-15116
; Sequence 15116, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15116
; LENGTH: 83462
; TYPE: DNA
; ORGANISM: Human
; FEATURE:

; NAME/KEY: misc feature
; LOCATION: (1)...(83462)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15116

Query Match 11.3%; Score 34; DB 3; Length 83462;
Best Local Similarity 54.9%; Pred. No. 7.6;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

56 TTTCAAAATAAACAGCTAACTCCAAATGTACCTAGGAGTTTATCATGAAGAAA 115
4080 TTTCAAAATGCTTCATTTCTCCAGTGATCTCCCTGGAGTGAGAAATGACTATGAAA 4139

116 TGTTAAATGCTTAAATTTTCAAAATAAGAGTAACCAAGAGTTTCAACATCTGTGTTAAAT 175
4140 TGTTAGATTTCTTTTCTTTTCTGTTGGGAATCAAGCAGCTTTTTCATGAGCAATTAATA 4199

176 AA 177
4200 AA 4201

RESULT 26

US-09-949-016-15638
; Sequence 15638, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15638
; LENGTH: 69764
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(69764)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15638

Query Match 11.3%; Score 33.8; DB 3; Length 69764;
Best Local Similarity 55.6%; Pred. No. 8.2; Mismatches 52; Indels 0; Gaps 0;

33 ATGAAGACTTGAATATTTCTTCAATTTGAAATAAACAGCTAACTCCAAATGTACCTAAC 92
29557 ATGAAGCTAGAACATTTCTAAATGTCCACAAAATGCTTTTATTAATTCGGTATAC 29616

93 TAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAAAATAAGAGTAAC 149
29617 CCCTGAAGTGAATCATTTAGGTAACTTTAGAGTTTAAATTTATGAAGTCTTTTAATAAC 29673

RESULT 27

US-09-543-681A-714/c
; Sequence 714, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: 2709.1002-001

| | | | |
|-------------------------------------------------------------------------------|-------|-------------------------------------------------------------|-------|
| DB | 395 | ACTATTAGAGAAATGATCATCATCTTCTGTAAGAAAAAATAAACTTCTAAGAGTATT | 454 |
| QY | 201 | GCTGTAGTAGAGTGCAATCTGTC | 222 |
| DB | 455 | GAGATGATTGAGTGAGTTGTGC | 476 |
| RESULT 31 | | | |
| US-09-949-016-12025 | | | |
| Sequence 12025, Application US/09949016 | | | |
| Patent No. 6812339 | | | |
| GENERAL INFORMATION: | | | |
| APPLICANT: VENTER, J. Craig et al. | | | |
| TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED | | | |
| TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF | | | |
| FILE REFERENCE: CL001307 | | | |
| CURRENT APPLICATION NUMBER: US/09/949,016 | | | |
| CURRENT FILING DATE: 2000-04-14 | | | |
| PRIOR APPLICATION NUMBER: 60/241,755 | | | |
| PRIOR FILING DATE: 2000-10-20 | | | |
| PRIOR APPLICATION NUMBER: 60/237,768 | | | |
| PRIOR FILING DATE: 2000-10-03 | | | |
| PRIOR APPLICATION NUMBER: 60/231,498 | | | |
| PRIOR FILING DATE: 2000-09-08 | | | |
| NUMBER OF SEQ ID NOS: 207012 | | | |
| SOFTWARE: FastSeq for Windows Version 4.0 | | | |
| SEQ ID NO 12025 | | | |
| LENGTH: 89268 | | | |
| TYPE: DNA | | | |
| ORGANISM: Human | | | |
| US-09-949-016-12025 | | | |
| Query Match 11.1%; Score 33.2; DB 3; Length 89268; | | | |
| Best Local Similarity 52.1%; Pred. No. 13; | | | |
| Matches 74; Conservative 0; Mismatches 68; Indels 0; Gaps 0; | | | |
| QY | 81 | ATTCTACTACTAGGGGAGTTTATCATGAAGAAATGTTTAAATGCTTTAAATTTTCAATA | 140 |
| DB | 30889 | ATTGTTCCAAAAGAGGACGAGGGTATCCAAACAGTTTCACCTTTTGAATGCTTAA | 30948 |
| QY | 141 | AGAGTAACACGAGTGTTGCAACATGCTGTTAAATAACCCGACAACTTCATCACTATA | 200 |
| DB | 30949 | ACTATTAGAGAAATGATCATCATCTTCTGTAAGAAAAAATAAACTTCTAAGAGTATT | 31008 |
| QY | 201 | GCTGTAGTAGAGTGCAATCTGTC | 222 |
| DB | 31009 | GAGATGATTGAGTGAGTTGTGC | 31030 |
| RESULT 32 | | | |
| US-09-949-016-13302 | | | |
| Sequence 13302, Application US/09949016 | | | |
| Patent No. 6812339 | | | |
| GENERAL INFORMATION: | | | |
| APPLICANT: VENTER, J. Craig et al. | | | |
| TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED | | | |
| TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF | | | |
| FILE REFERENCE: CL001307 | | | |
| CURRENT APPLICATION NUMBER: US/09/949,016 | | | |
| CURRENT FILING DATE: 2000-04-14 | | | |
| PRIOR APPLICATION NUMBER: 60/241,755 | | | |
| PRIOR FILING DATE: 2000-10-20 | | | |
| PRIOR APPLICATION NUMBER: 60/237,768 | | | |
| PRIOR FILING DATE: 2000-10-03 | | | |
| PRIOR APPLICATION NUMBER: 60/231,498 | | | |
| PRIOR FILING DATE: 2000-09-08 | | | |
| NUMBER OF SEQ ID NOS: 207012 | | | |
| SOFTWARE: FastSeq for Windows Version 4.0 | | | |
| SEQ ID NO 13302 | | | |
| LENGTH: 90544 | | | |
| TYPE: DNA | | | |
| ORGANISM: Human | | | |

US-09-949-016-13302

Query Match 11.1%; Score 33.2; DB 3; Length 90544;
Best Local Similarity 52.1%; Pred. No. 13;
Matches 74; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 81 ATTGTACCTTAACCTAGGGGAGTTTATCATGACGAAGATGTTTAAATGCTTTAAATTTTCAATA 140
DB 31074 ATTGTTTACAAAGAGACGAGGAGGTAATCCAAACAGTTTCTACTTTTGTGATTGCTTAACA 31133
QY 141 AGAAGTAACACAGAGTGTTCGAACATGCTGTTAAATTAACCCGACAACTTCAATCACTATA 200
DB 31134 ACTATTAGAAATATGATCATCTTCTGTGAAGAAAAAATAAATCTTCTAGAGTATT 31193
QY 201 GCTGTAGTAGTCATTCTGC 222
DB 31194 GAGATGATTGAGTGAGTTGTGC 31215

RESULT 33
US-09-621-976-2813/c
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match 11.0%; Score 33; DB 3; Length 832;
Best Local Similarity 10.3%; Pred. No. 3.6;
Matches 18; Conservative 91; Mismatches 66; Indels 0; Gaps 0;

QY 26 GACTACGATCAAGACTTGAAATATCTTTCATTTGAAATAAACAGCTAACTCCCAATTGT 85
DB 194 SNNYTCACWGRWSTYTMANGKKWRYATTFRSANNMMAAAATMMWMMWMAWMSRSG 135
QY 86 ACCTAACTAGGGGAGTTTATCATGAAGAAATGTTTAAATGCTTTAAATTTTCAATAAAGAAG 145
DB 134 AAMYRTMMWGYRYWRKKSRYRTRCAWAYAKTKRSYYCWRKWKRCMMMMMAWAY 75
QY 146 TAACACGAGTTGTGCAACATGCTGTTAAATAACCCGACAACTTCAATCACTATA 200
DB 74 GKTMMRACWKTRYERWAWAWRMMWTTMMMYYYWTRAMKRWWRKWSNM 20

RESULT 34
US-08-929-329-3/c
; Sequence 3, Application US/08929329
; Patent No. 612070
; GENERAL INFORMATION:
; APPLICANT: Adams, John H
; APPLICANT: Dalton, John P
; APPLICANT: Kappe, Stefan
; TITLE OF INVENTION: Plasmodium Proteins Useful for Preparing
; TITLE OF INVENTION: Vaccine Compositions
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESS: Barnes & Thornburg
; STREET: 11 S Meridian

CITY: Indianapolis
 STATE: Indiana
 COUNTRY: USA
 ZIP: 46204
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/929,329
 FILING DATE:
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Breen, John P
 REGISTRATION NUMBER: 38,933
 REFERENCE/DOCKET NUMBER: 835910-28685
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (317) 231-7745
 TELEFAX: (317) 231-7433
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1194 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Plamodium berghei
 US-08-929-329-3
 Query Match 11.0%; Score 33; DB 3; Length 1194;
 Best Local Similarity 48.2%; Pred. No. 4;
 Matches 93; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
 QY 49 TTCTTCATTGGAATAAAGAGCTAACTCCCAATTTGTACCTAACCTAGGGAGTTTATCAT 108
 DB 1083 TTGTTGTTAGCTAAAGAGCATGAAATAAATAATATATATACGTATATTTTCAA 1024
 QY 109 GAAGAAATGTTAAATGCTTAATTTTCAAATGAAGTAACAGAGTGTTCGAACATGCT 168
 DB 1023 TAGCTAGTATTTAAATGAATAAATTTTAAATATAAAAAAATAACAAATTTAAATATT 964
 QY 169 GTTAAATAACCCGACAACTTCAATCACTATAGCTGTAGTAGTGCTTCGCAAGGAT 228
 DB 963 TTTTGAATATTAATTAACATATACATATAGCTTAACGTATAGACTTTATCATAACTTCAACAT 904
 QY 229 CCCAGAGTAACCA 241
 DB 903 TTCATCAAAAGCA 891
 RESULT 35
 US-09-662-254B-24/c
 Sequence 24, Application US/09662254B
 Patent No. 6933145
 GENERAL INFORMATION:
 APPLICANT: Moyer, Richard W.
 APPLICANT: Li, Yi
 APPLICANT: Bowden, Alison Louise
 TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous
 TITLE OF INVENTION: Vertebrate Cells
 FILE REFERENCE: UF-221C1XC1
 CURRENT APPLICATION NUMBER: US/09/662,254B
 CURRENT FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 09/086,651
 PRIOR FILING DATE: 1998-05-29
 PRIOR APPLICATION NUMBER: 60/224,479
 PRIOR FILING DATE: 2000-08-10
 NUMBER OF SEQ ID NOS: 80
 SOFTWARE: Patent In version 3.1

```

; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13188
; LENGTH: 266748
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-13188

Query Match          11.0%; Score 33; DB 3; Length 266748;
Best Local Similarity 57.1%; Pred. No. 22;
Matches 60; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 161 AACATGCTGTTAAATAACCCGACAAACTTCAATCACTATAGCTGTAGTAGAGTGCAATTCT 220
DB 171679 AACATCCTTTTCATATATTTCTCGCGCACTTCTATCAGTATATATGCAGGATAGAGTACTAG 171620

QY 221 GCAAGGATCCGAGGTAAACCATGATTTTGGAAATGCAATGTTGAA 265
DB 171619 ACATGGACCTGGGAGGTCAAGAATAGGTGTATTTTAAATTTTAA 171575

RESULT 40
US-09-949-016-92131/c
; Sequence 92131, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 92131
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-92131

Query Match          10.9%; Score 32.8; DB 3; Length 601;
Best Local Similarity 49.1%; Pred. No. 3.7;
Matches 114; Conservative 0; Mismatches 117; Indels 1; Gaps 1;

QY 30 ACAGTGAAGACTTGAATAATTTCTTCATTTGAAATAAACAGCTAACTCCCAAAATGTGACTT 89
DB 308 ACAATCASCATCTTGACTTTTACTTCTCA-GGAAAATAAAGTTTACTTGGAAACAGATATTT 250

QY 90 AACTAGGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAATAAAGAGTAAC 149
DB 249 TAGCATAAGAACCTTAGCATGAAGCAAACTAAATAAGTATAGTATTCAAAGGTCACAAAT 190

QY 150 CAGAGTGTTTGCAACATGCTGTGTTAAATAACCCGACAAACTTCAATCACTATAGCTGTAGTA 209
DB 189 CATATTGGCCAAATCCCTTCAAAGCCAGGTACCACTACTACAGACAGATCTGGGTGA 130

QY 210 GAGTGCATTTCTGCAAGGATCCCGAGTAAACGAGTATTTTGGAAATGCAATCT 261
DB 129 CGATGCGTCTTATTTGCAACAGTGAACCTTCTGGACCTTAGAATCTGCGCCTCTG 78

```

```

RESULT 41
US-09-949-016-16290
; Sequence 16290, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16290
; LENGTH: 10783
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16290

Query Match 10.9%; Score 32.8; DB 3; Length 10783;
Best Local Similarity 56.5%; Pred. No. 9.1;
Matches 61; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 27 ACTACGATGAAGACTTGAATATTTCTTCATTTGAAATAAACAGCTAACTCCCAATTGTA 86
DB 7795 ACAACAGTCAATAAGAGAAAGTACAAAATTCAAAGTAAGAAATTAACCTTGGATATTGAA 7854

QY 87 CCTACTAGGGAGCTTTATCATGAGAAATGTTTAAATGCTTAATTTT 134
DB 7855 GCTTACTTTTCTAGATTATGTTTCAGAGATCTTTGACATTTTAACTTT 7902

RESULT 42
US-09-949-016-13659
; Sequence 13659, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13659
; LENGTH: 23105
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13659

Query Match 10.9%; Score 32.8; DB 3; Length 23105;
Best Local Similarity 49.4%; Pred. No. 12;
Matches 85; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 129 AATTTCATAAAGAGTAACAGAGTGTTCGAACATGCTGTTAAATAACCCGACAACT 188
DB 7977 AATTTCATTGATGGAATTACCAATAGTTTGCTAAATACATTTTAATAAGTCCCTAGACT 8036

```

```

QY 189 TCAATCACTATAGCTGTAGTAGTGCAATTCCTCAAGGATCCAGAGTAACCAAGTATTTT 248
DB 8037 TTGAACCTCTTCCCTCTGGAGCAGACCTTGGGGCAAGGATGTGGGAGCATGTAGTTATT 8096

QY 249 GGAATGCAATGTTGAACCGACCATACTAATTAATCTGCTGATTAGGAAGTAT 300
DB 8097 TGGAGGTGATCTTAGGAAAGAGGGATTAAGGAGGAACCTTAGGAGGACTAT 8148

RESULT 43
US-09-949-016-12582
; Sequence 12582, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12582
; LENGTH: 23108
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12582

Query Match 10.9%; Score 32.8; DB 3; Length 23108;
Best Local Similarity 49.4%; Pred. No. 12;
Matches 85; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 129 AATTTCATAAAGAGTAACAGAGTGTTCGAACATGCTGTTAAATAACCCGACAACT 188
DB 7980 AATTTCATTGATGGAATTTACCATAGTTTGCTAAATACATTTTAATAAGTCCCTAGACT 8039

QY 189 TCAATCACTATAGCTGTAGTAGTGCAATTCCTCAAGGATCCAGAGTAACCAAGTATTTT 248
DB 8040 TTGAACCTCTTCCCTCTGGAGCAGACCTTGGGGCAAGGATGTGGGAGCATGTAGTTATT 8099

QY 249 GGAATGCAATGTTGAACCGACCATACTAATTAATCTGCTGATTAGGAAGTAT 300
DB 8100 TGGAGGTGATCTTAGGAAAGAGGGATTAAGGAGGAACCTTAGGAGGACTAT 8151

RESULT 44
US-09-949-016-13176
; Sequence 13176, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13176
; LENGTH: 36577
; TYPE: DNA

```

```

; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(36577)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13176

Query Match      10.9%; Score 32.8; DB 3; Length 36577;
Best Local Similarity 50.6%; Pred. No. 13;
Matches 79; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 33 ATGAAGACTTGAAATATTTCTTCATTTGAAATAAAGACAGCTAACTCCCAATTTGTACTTAAC 92
DB 21634 AGGAAGATGCATATTCCTGCTTTTGGGAATAAGAACCAACTGGAAATCCAGGTGGC 21693

QY 93 TAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTTAAATTTTCAAAATGAAGATTAACCGAG 152
DB 21694 TAGTGGTGTAGACTGTTAAGTAATGTTTAAAGATGTGAGTTTAGCATAATAGATTGAAT 21753

QY 153 AGTGTGCAACATGCTGTTAAATAACCCGACAAACT 188
DB 21754 TGATTAGATACAAATGGTTTAAATAATTTTAAAAAAT 21789

RESULT 45
US-09-596-002-40 :
; Sequence 40, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 40
; LENGTH: 119211
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte template ID No. 6632636 40
; PUBLICATION INFORMATION:
US-09-596-002-40

Query Match      10.9%; Score 32.8; DB 3; Length 119211;
Best Local Similarity 52.1%; Pred. No. 19;
Matches 73; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 1 GGCATTGACGGCTTAATGCTTAAGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
DB 108036 GGGCTGGATAAAGTGAAGTGGTGCAGCAATCAATCAATCAATCAATCAATCAATCAATCAAT 108095

QY 61 AATAAAGCAGCTAACTCCCAATTTGTACCTACTAGGGGAGTTTATCATGAGAGAAATGTTT 120
DB 108096 AATCAAAATTTTATTCCTCAATATGCGATCAATATTAATAATTTATGTCATATCAAAAGTTT 108155

QY 121 AAATGCTTTAATTTTCAATA 140
DB 108156 TAAAGTGATATTTCAATA 108175

RESULT 46
US-09-949-016-14382/c
; Sequence 14382, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

```

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14382
; LENGTH: 192956
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(192956)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14382

Query Match      10.9%; Score 32.8; DB 3; Length 192956;
Best Local Similarity 49.1%; Pred. No. 22;
Matches 114; Conservative 0; Mismatches 117; Indels 1; Gaps 1;

QY 30 ACGATGAAGACTTGAAATATTTCTTCATTTGAAATAAAGACAGCTAACTCCCAATTTGTACTT 89
DB 4134 ACAATCAGCACTTGACTTTTACTTCA-GGAAATAAAGTTTACTTTGGGAACAGATATTT 4076

QY 90 AACTAGGGGAGCTTATCATGAAGAAATGTTTAAATGCTTTAAATTTTCAAAATGAAGATTAAC 149
DB 4075 TAGCATGAAGACTTAGCATAAGAACAAACTAAATGATTAGTTATTTCCAAAGGTACAAAAT 4016

QY 150 CAGAGTGTGCAACATGCTGTTAAATAACCCGACAACTTCAATCACTATAGCTGTAGTA 209
DB 4015 CATATTTGGCCAAAATCCTTCAAAAGCCAAAGTACCCTACTACTACAGACAGATCTGGTGA 3956

QY 210 GAGTGCATTTCTGCAAGGATCCAGAGTAACAGATTTTGGAAATGCAATGT 261
DB 3955 CGATGCGTCTTATTGCAACAGTGAACCTTGGACCTTAGAATGTGCCCTGT 3904

RESULT 47
US-10-002-623-923/c
; Sequence 923, Application US/10002623
; Patent No. 6929911
; GENERAL INFORMATION:
; APPLICANT: OEFNER, PETER J.
; APPLICANT: UNDERHILL, PETER A.
; TITLE OF INVENTION: A METHOD FOR DETERMINING GENETIC
; AFFILIATION, SUBSTRUCTURE AND GENE FLOW WITHIN HUMAN
; TITLE OF INVENTION: POPULATIONS
; FILE REFERENCE: STAN-212
; CURRENT APPLICATION NUMBER: US/10/002,623
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/245,355
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 952
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 923
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-002-623-923

Query Match      10.9%; Score 32.6; DB 3; Length 456;
Best Local Similarity 60.9%; Pred. No. 3.9;
Matches 53; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 55 ATTTGAAATAAAGCTTAACCTCCCAATTTGTACTTAACCTAGGGAGTTTATCATCAAGAA 114

```

Db 297 ATTCCAAAGAAAAAATAAATTTTGTAAATATAGACTTCTTATGTTGTAATA 238
Qy 115 ATGTTTAAATGCTTAATTTTCAATAA 141
Db 237 ATTTATAATGCATATATACAATACAA 211

RESULT 48

US-10-002-623-926/c
; Sequence 926, Application US/10002623
; Patent No. 6929911
; GENERAL INFORMATION:
; APPLICANT: OFFNER, PETER J.
; APPLICANT: UNDERHILL, PETER A.
; TITLE OF INVENTION: A METHOD FOR DETERMINING GENETIC
; TITLE OF INVENTION: AFFILIATION, SUBSTRUCTURE AND GENE FLOW WITHIN HUMAN
; TITLE OF INVENTION: POPULATIONS
; FILE REFERENCE: STAN-212
; CURRENT APPLICATION NUMBER: US/10/002,623
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/245,355
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 952
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 926
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-002-623-926

Query Match 10.9%; Score 32.6; DB 3; Length 456;
Best Local Similarity 60.9%; Pred. No. 3.9;
Matches 53; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 55 ATTGGAATAACAGCTAACTCCCAATTTGTACCTACTAGGGGAGTTTATCATGAAGAA 114
Db 297 ATTCCAAAGAAAAAATAAATTTTGTAAATATAGACTTCTTATGTTGTAATA 238

Qy 115 ATGTTTAAATGCTTAATTTTCAATAA 141
Db 237 ATTTATAATGCATATATACAATACAA 211

RESULT 49

US-10-002-623-929/c
; Sequence 929, Application US/10002623
; Patent No. 6929911
; GENERAL INFORMATION:
; APPLICANT: OFFNER, PETER J.
; APPLICANT: UNDERHILL, PETER A.
; TITLE OF INVENTION: A METHOD FOR DETERMINING GENETIC
; TITLE OF INVENTION: AFFILIATION, SUBSTRUCTURE AND GENE FLOW WITHIN HUMAN
; TITLE OF INVENTION: POPULATIONS
; FILE REFERENCE: STAN-212
; CURRENT APPLICATION NUMBER: US/10/002,623
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/245,355
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 952
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 929
; LENGTH: 623
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-002-623-929

Query Match 10.9%; Score 32.6; DB 3; Length 623;
Best Local Similarity 60.9%; Pred. No. 4.3;
Matches 53; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 55 ATTGGAATAACAGCTAACTCCCAATTTGTACCTACTAGGGGAGTTTATCATGAAGAA 114
Db 297 ATTCCAAAGAAAAAATAAATTTTGTAAATATAGACTTCTTATGTTGTAATA 238

Qy 115 ATGTTTAAATGCTTAATTTTCAATAA 141
Db 237 ATTTATAATGCATATATACAATACAA 211

RESULT 50

US-09-809-545A-64/c
; Sequence 64, Application US/09809545A
; Patent No. 6800455
; GENERAL INFORMATION:
; APPLICANT: Stanton, Lawrence W.
; APPLICANT: White, R. Tyler
; TITLE OF INVENTION: SECRETED FACTORS
; FILE REFERENCE: SCIOS.017A
; CURRENT APPLICATION NUMBER: US/09/809,545A
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 819
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (0)...(0)
; OTHER INFORMATION: n = A, T, C, or G
US-09-809-545A-64

Query Match 10.9%; Score 32.6; DB 3; Length 819;
Best Local Similarity 47.0%; Pred. No. 4.7;
Matches 101; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Qy 39 ACTTGAATAATCTTCTTCAATTTGAAATAAACAAGCTAACTCCCAATTTGTACCTAACTAGGG 98
Db 371 ACTTGAATAACCACTTTTAAAAAGGCAACAAGTGACCAGATTTAAATTAAGATAGT 312

Qy 99 AGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAATAAGAAGTAACCAGAGTGT 158
Db 311 ATATTGCAAAATACATCAGCATATATATGTCGAGCACAAGCAAGTAGTTTCCATTG 252

Qy 159 GCAACATGCTTTAAATAACCCGACAACTTCAATCACTATAGCTAGTAGTAGTGCATT 218
Db 251 GAAGGATACCTTTTAACTTCTAATAAGCAACCCTTCCAAATTTGTGTGGTAGTTCAGT 192

Qy 219 CTGCAAGGATCCAGAGTAACCCAGTATTTTGGAAA 253
Db 191 CTAATATATACAAGAGTGGATACATTTTACACA 157

Search completed: July 18, 2006, 01:05:41
Job time : 140 secs

GenCore version 5.1.9
 Copyright (c) 1993 - 2006 Bioceleration Ltd.
 OM nucleic - nucleic search, using sw model
 Run on: July 17, 2006, 22:44:13 ; Search time 1462 Seconds
 (without alignments)
 2521.403 Million cell updates/sec

Title: SEQ1-96535C
 Perfect score: 300
 Sequence: 1 ggcattgacggctaaatgc.....atctgctgattaggaagtat 300

Scoring table: IDENTITY NUC
 Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA Main.*

1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq.*
 2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*
 3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
 4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
 5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
 6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
 7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
 8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
 9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
 10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
 11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
 12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
 13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
 14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
 15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
 16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|-------|-------------|-----------|----|-----------------------|
| 1 | 299.6 | 100.0 | 447894 | 10 | US-10-723-670-1 |
| 2 | 298 | 99.3 | 614 | 4 | US-09-925-065A-428890 |
| 3 | 298 | 99.3 | 614 | 5 | US-09-925-065A-428890 |
| 4 | 222 | 74.0 | 627 | 6 | US-10-027-632-116164 |
| 5 | 222 | 74.0 | 627 | 6 | US-10-027-632-116165 |
| 6 | 222 | 74.0 | 627 | 7 | US-10-027-632-116164 |
| 7 | 222 | 74.0 | 627 | 7 | US-10-027-632-116165 |
| 8 | 40.6 | 13.5 | 201 | 10 | US-10-995-561-33431 |
| 9 | 40.6 | 13.5 | 201 | 10 | US-10-995-561-37932 |
| 10 | 40.6 | 13.5 | 588 | 4 | US-09-925-065A-349719 |
| 11 | 40.6 | 13.5 | 588 | 5 | US-09-925-065A-349719 |
| 12 | 40.6 | 13.5 | 599 | 12 | US-10-301-480-421779 |
| 13 | 40.6 | 13.5 | 599 | 12 | US-10-301-480-1035188 |
| 14 | 40.6 | 13.5 | 57198 | 10 | US-10-995-561-13263 |
| 15 | 40.6 | 13.5 | 173995 | 10 | US-10-995-561-13277 |
| 16 | 39.6 | 13.2 | 483 | 12 | US-10-301-480-21248 |
| 17 | 39.6 | 13.2 | 483 | 12 | US-10-301-480-634657 |

| | | | | | | |
|----|------|------|--------|----|-----------------------|---------------------|
| 18 | 39.6 | 13.2 | 502 | 6 | US-10-027-632-93552 | Sequence 93552, A |
| 19 | 39.6 | 13.2 | 502 | 6 | US-10-027-632-305099 | Sequence 305099, A |
| 20 | 39.6 | 13.2 | 502 | 7 | US-10-027-632-93552 | Sequence 93552, A |
| 21 | 39.6 | 13.2 | 502 | 7 | US-10-027-632-305099 | Sequence 305099, A |
| 22 | 39.4 | 13.1 | 588 | 4 | US-09-925-065A-349720 | Sequence 349720, A |
| 23 | 39.4 | 13.1 | 588 | 4 | US-09-925-065A-349722 | Sequence 349722, A |
| 24 | 39.4 | 13.1 | 588 | 5 | US-09-925-065A-349720 | Sequence 349720, A |
| 25 | 39.4 | 13.1 | 588 | 5 | US-09-925-065A-349722 | Sequence 349722, A |
| 26 | 39.4 | 13.1 | 599 | 12 | US-10-301-480-421780 | Sequence 421780, A |
| 27 | 39.4 | 13.1 | 599 | 12 | US-10-301-480-421782 | Sequence 421782, A |
| 28 | 39.4 | 13.1 | 599 | 12 | US-10-301-480-1035189 | Sequence 1035189, A |
| 29 | 39.4 | 13.1 | 599 | 12 | US-10-301-480-1035191 | Sequence 1035191, A |
| 30 | 39.4 | 13.1 | 602 | 4 | US-09-925-065A-349721 | Sequence 349721, A |
| 31 | 39.4 | 13.1 | 602 | 5 | US-09-925-065A-349721 | Sequence 349721, A |
| 32 | 39.4 | 13.1 | 608 | 12 | US-10-301-480-421781 | Sequence 421781, A |
| 33 | 39.4 | 13.1 | 608 | 12 | US-10-301-480-421782 | Sequence 421782, A |
| 34 | 39.4 | 13.1 | 60461 | 7 | US-10-301-480-1035190 | Sequence 1035190, A |
| 35 | 38.8 | 12.9 | 295096 | 6 | US-10-087-132-331 | Sequence 82, Appl |
| 36 | 38.6 | 12.9 | 2224 | 9 | US-10-357-930-21832 | Sequence 331, Appl |
| 37 | 38.6 | 12.9 | 2224 | 9 | US-10-357-930-22160 | Sequence 21832, A |
| 38 | 38.6 | 12.9 | 2224 | 9 | US-10-357-930-22166 | Sequence 22160, A |
| 39 | 38.6 | 12.9 | 2224 | 9 | US-10-357-930-23173 | Sequence 22166, A |
| 40 | 38.6 | 12.9 | 2224 | 9 | US-10-357-930-24502 | Sequence 23173, A |
| 41 | 38.6 | 12.9 | 2224 | 9 | US-10-357-930-24579 | Sequence 24502, A |
| 42 | 38.6 | 12.9 | 2224 | 9 | US-10-357-930-24590 | Sequence 24579, A |
| 43 | 38.6 | 12.9 | 2224 | 9 | US-10-357-930-25046 | Sequence 24590, A |
| 44 | 38.6 | 12.9 | 2224 | 9 | US-10-357-930-27679 | Sequence 25046, A |
| 45 | 38.6 | 12.9 | 2224 | 9 | US-10-357-930-28021 | Sequence 27679, A |
| 46 | 38.6 | 12.9 | 2224 | 9 | US-10-357-930-28027 | Sequence 28021, A |
| 47 | 38.6 | 12.9 | 2224 | 9 | US-10-357-930-29041 | Sequence 28027, A |
| 48 | 38.2 | 12.7 | 552 | 6 | US-10-027-632-86622 | Sequence 29041, A |
| 49 | 38.2 | 12.7 | 552 | 6 | US-10-027-632-86623 | Sequence 86622, A |
| 50 | 38.2 | 12.7 | 552 | 6 | US-10-027-632-316419 | Sequence 86623, A |
| 51 | 38.2 | 12.7 | 552 | 6 | US-10-027-632-316420 | Sequence 316419, A |
| 52 | 38.2 | 12.7 | 552 | 7 | US-10-027-632-86622 | Sequence 316420, A |
| 53 | 38.2 | 12.7 | 552 | 7 | US-10-027-632-86623 | Sequence 86622, A |
| 54 | 38.2 | 12.7 | 552 | 7 | US-10-027-632-316419 | Sequence 86623, A |
| 55 | 38.2 | 12.7 | 552 | 7 | US-10-027-632-316420 | Sequence 316419, A |
| 56 | 37.8 | 12.6 | 537 | 12 | US-10-301-480-18060 | Sequence 18060, A |
| 57 | 37.8 | 12.6 | 537 | 12 | US-10-301-480-631469 | Sequence 631469, A |
| 58 | 37.6 | 12.5 | 614 | 12 | US-10-301-480-248299 | Sequence 248299, A |
| 59 | 37.6 | 12.5 | 614 | 12 | US-10-301-480-248300 | Sequence 248300, A |
| 60 | 37.6 | 12.5 | 614 | 12 | US-10-301-480-861708 | Sequence 861708, A |
| 61 | 37.6 | 12.5 | 614 | 12 | US-10-301-480-861709 | Sequence 861709, A |
| 62 | 37.6 | 12.5 | 617 | 4 | US-09-925-065A-154831 | Sequence 861709, A |
| 63 | 37.6 | 12.5 | 617 | 4 | US-09-925-065A-154832 | Sequence 154831, A |
| 64 | 37.6 | 12.5 | 617 | 5 | US-09-925-065A-154831 | Sequence 154832, A |
| 65 | 37.6 | 12.5 | 617 | 5 | US-09-925-065A-154832 | Sequence 154831, A |
| 66 | 37.6 | 12.5 | 625 | 12 | US-10-301-480-203218 | Sequence 154832, A |
| 67 | 37.6 | 12.5 | 625 | 12 | US-10-301-480-816627 | Sequence 203218, A |
| 68 | 37.6 | 12.5 | 626 | 4 | US-09-925-065A-102887 | Sequence 816627, A |
| 69 | 37.6 | 12.5 | 626 | 5 | US-09-925-065A-102887 | Sequence 102887, A |
| 70 | 37.4 | 12.5 | 537 | 12 | US-10-301-480-18061 | Sequence 102887, A |
| 71 | 37.4 | 12.5 | 537 | 12 | US-10-301-480-631470 | Sequence 18061, A |
| 72 | 37.4 | 12.5 | 635 | 4 | US-09-925-065A-929082 | Sequence 631470, A |
| 73 | 37.4 | 12.5 | 635 | 4 | US-09-925-065A-929082 | Sequence 929082, A |
| 74 | 37.4 | 12.5 | 635 | 5 | US-09-925-065A-950049 | Sequence 950049, A |
| 75 | 37.4 | 12.5 | 635 | 5 | US-09-925-065A-929082 | Sequence 929082, A |
| 76 | 37 | 12.3 | 6857 | 8 | US-10-455-822-16 | Sequence 950049, A |
| 77 | 37 | 12.3 | 6857 | 8 | US-10-455-822-50 | Sequence 16, Appl |
| 78 | 37 | 12.3 | 6857 | 8 | US-10-455-822-52 | Sequence 50, Appl |
| 79 | 37 | 12.3 | 6857 | 8 | US-10-455-822-54 | Sequence 52, Appl |
| 80 | 37 | 12.3 | 6857 | 8 | US-10-455-822-56 | Sequence 54, Appl |
| 81 | 37 | 12.3 | 6857 | 8 | US-10-455-822-58 | Sequence 56, Appl |
| 82 | 37 | 12.3 | 6857 | 8 | US-10-455-822-60 | Sequence 58, Appl |
| 83 | 37 | 12.3 | 6857 | 8 | US-10-455-822-62 | Sequence 60, Appl |
| 84 | 37 | 12.3 | 6857 | 8 | US-10-455-822-64 | Sequence 62, Appl |
| 85 | 37 | 12.3 | 6857 | 8 | US-10-455-822-66 | Sequence 64, Appl |
| 86 | 37 | 12.3 | 6857 | 8 | US-10-455-822-68 | Sequence 66, Appl |
| 87 | 37 | 12.3 | 6857 | 8 | US-10-455-822-70 | Sequence 68, Appl |
| 88 | 37 | 12.3 | 6857 | 8 | US-10-455-822-72 | Sequence 70, Appl |
| 89 | 37 | 12.3 | 6857 | 8 | US-10-455-822-74 | Sequence 72, Appl |
| 90 | 37 | 12.3 | 6857 | 8 | US-10-455-822-76 | Sequence 74, Appl |
| 91 | 37 | 12.3 | 6857 | 8 | US-10-455-822-78 | Sequence 76, Appl |

```

37 12.3 6857 8 US-10-455-822-187
37 12.3 6857 8 US-10-455-822-189
37 12.3 6857 8 US-10-407-484-16
37 12.3 6857 8 US-10-407-484-50
37 12.3 6857 8 US-10-407-484-52
37 12.3 6857 8 US-10-407-484-54
37 12.3 6857 8 US-10-407-484-56
37 12.3 6857 8 US-10-407-484-58
37 12.3 6857 8 US-10-407-484-60
37 12.3 6857 8 US-10-407-484-62
37 12.3 6857 8 US-10-407-484-64
37 12.3 6857 8 US-10-407-484-66
37 12.3 6857 8 US-10-407-484-68
37 12.3 6857 8 US-10-407-484-70
37 12.3 6857 8 US-10-407-484-72
37 12.3 6857 8 US-10-407-484-74
37 12.3 6857 8 US-10-407-484-76
37 12.3 6857 8 US-10-407-484-187
37 12.3 6857 8 US-10-407-484-189
37 12.3 6857 16 US-11-068-859-16
37 12.3 6857 16 US-11-068-859-50
37 12.3 6857 16 US-11-068-859-52
37 12.3 6857 16 US-11-068-859-54
37 12.3 6857 16 US-11-068-859-56
37 12.3 6857 16 US-11-068-859-58
37 12.3 6857 16 US-11-068-859-60
37 12.3 6857 16 US-11-068-859-62
37 12.3 6857 16 US-11-068-859-64
37 12.3 6857 16 US-11-068-859-66
37 12.3 6857 16 US-11-068-859-68
37 12.3 6857 16 US-11-068-859-70
37 12.3 6857 16 US-11-068-859-72
37 12.3 6857 16 US-11-068-859-74
37 12.3 6857 16 US-11-068-859-76
37 12.3 6857 16 US-11-068-859-187
37 12.3 6857 16 US-11-068-859-189
36.6 12.2 578 4 US-09-925-065A-451737
36.6 12.2 578 5 US-09-925-065A-451737
36.6 12.2 578 12 US-10-301-480-510754
36.6 12.2 578 12 US-10-301-480-1124163
36.2 12.1 707 8 US-10-424-599-53104
36.2 12.1 9087 8 US-10-221-613-140
36.2 12.1 17738 7 US-10-311-455-1511
36.2 12.1 37992 8 US-10-322-281-779
36 12.0 6020 7 US-10-311-455-172
36 12.0 6020 10 US-10-240-708-8
36 12.0 598 4 US-09-925-065A-216848
35.8 11.9 598 5 US-09-925-065A-216848
35.8 11.9 600 12 US-10-301-480-303243
35.8 11.9 600 12 US-10-301-480-916652
35.8 11.9 2310 9 US-10-473-126-274
35.8 11.9 7588 6 US-10-239-676-191
35.8 11.9 7588 7 US-10-311-455-1985
35.8 11.9 7588 7 US-10-240-453-283
35.6 11.9 277616 8 US-10-367-094-83
35.6 11.8 614 12 US-10-301-480-333833
35.4 11.8 614 12 US-10-301-480-947242
35.4 11.8 618 4 US-09-925-065A-254972
35.4 11.8 618 5 US-09-925-065A-254972
35.4 11.8 999 12 US-10-301-480-541365

```

ALIGNMENTS

```

RESULT 1
US-10-723-670-1
; Sequence 1, Application US/10723670
; Publication No. US20050118606A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, RICHARD B.
; APPLICANT: NELSON, MATTHEW ROBERTS
; APPLICANT: KAMMERER, STEFAN M.

```

```

; APPLICANT: BRAUN, ANDREAS
; APPLICANT: RENEHARD, RIKARD
; TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF BREAST CANCER AND TREATMENTS
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: SEQ-4067-UT
; CURRENT APPLICATION NUMBER: US/10/723,670
; CURRENT FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: 60/429,136
; PRIOR FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 60/490,234
; PRIOR FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: 60/467,823
; PRIOR FILING DATE: 2003-05-02
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1
; TYPE: DNA
; LENGTH: 447894
; ORGANISM: Homo sapiens
; US-10-723-670-1

Query Match 100.0%; Score 300; DB 10; Length 447894;
Best Local Similarity 99.7%; Pred. No. 1.3e-69;
Matches 299; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATTGACAGGCTAAATGCTAAGTACGACTACGATGAGACTTGAATATTTCTTCATTGA 60
Db 96401 GGCATTGACAGGCTAAATGCTAAGTACGACTACGATGAGACTTGAATATTTCTTCATTGA 96460

Qy 61 AATAACAGCTAACTCCCAATTTGTACCTAACTAGGAGGAGTTTATCATGAAGAATGTTT 120
Db 96461 AATAACAGCTAACTCCCAATTTGTACCTAACTAGGAGGAGTTTATCATGAAGAATGTTT 96520

Qy 121 AATGCTTAAATTTTCAAAATGAAGTAACAGAGTGTTGCAACATGCTGTTTAAATAACCC 180
Db 96521 AATGCTTAAATTTTCAAAATGAAGTAACAGAGTGTTGCAACATGCTGTTTAAATAACCC 96580

Qy 181 GACAACTTCAATCATTATAGCTGTAGTAGAGTGATTCGACAGGATCCAGAGTAACC 240
Db 96581 GACAACTTCAATCATTATAGCTGTAGTAGAGTGATTCGACAGGATCCAGAGTAACC 96640

Qy 241 AGTATTTTGAATGCAATGTTGAACCGACCATCTACTAATATCTGCTGATTTAGGAAGTAT 300
Db 96641 AGTATTTTGAATGCAATGTTGAACCGACCATCTACTAATATCTGCTGATTTAGGAAGTAT 96700

RESULT 2
US-09-925-065A-428890
; Sequence 428890, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428890
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapiens

```



```
US-09-925-065A-428890
Query Match          99.3%; Score 298; DB 4; Length 614;
Best Local Similarity 99.3%; Pred. No. 2.2e-70;
Matches 298; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGCATTGACAGGCTAAATGCTAAGTACTAGTACGATGAGACTTGAATATCTTCATTGGA 60
Db 105 GGCATTGACAGGCTAAATGCTAAGTACTAGTACGATGAGACTTGAATATCTTCATTGGA 164

Qy 61 AATAAAGCTTAATTTTCAATTAAGAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 120
Db 165 AATAAAGCTTAATTTTCAATTAAGAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 224

Qy 121 AAATGCTTAATTTTCAATTAAGAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 180
Db 225 AAATGCTTAATTTTCAATTAAGAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 284

Qy 181 GACAACTTCAATCACTATAGTCTGTAGTAGTGCATCTGCAAGGATCCAGAGTAACC 240
Db 285 GACAACTTCAATCACTATAGTCTGTAGTAGTGCATCTGCAAGGATCCAGAGTAACC 344

Qy 241 AGTATTTTGGAAATGCAATGTTGAACCCAGCATCTAATTAATCTGCTGATTAGGAAGTAT 300
Db 345 AGTATTTTGGAAATGCAATGTTGAACCCAGCATCTAATTAATCTGCTGATTAGGAAGTAT 404

RESULT 3
US-09-925-065A-428890
; Sequence 428890; Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428890
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-428890

Query Match          99.3%; Score 298; DB 5; Length 614;
Best Local Similarity 99.3%; Pred. No. 2.2e-70;
Matches 298; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGCATTGACAGGCTAAATGCTAAGTACTAGTACGATGAGACTTGAATATCTTCATTGGA 60
Db 105 GGCATTGACAGGCTAAATGCTAAGTACTAGTACGATGAGACTTGAATATCTTCATTGGA 164

Qy 61 AATAAAGCTTAATTTTCAATTAAGAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 120
Db 165 AATAAAGCTTAATTTTCAATTAAGAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 224

Qy 121 AAATGCTTAATTTTCAATTAAGAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 180
Db 225 AAATGCTTAATTTTCAATTAAGAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 284

Qy 181 GACAACTTCAATCACTATAGTCTGTAGTAGTGCATCTGCAAGGATCCAGAGTAACC 240
Db 285 GACAACTTCAATCACTATAGTCTGTAGTAGTGCATCTGCAAGGATCCAGAGTAACC 344

US-09-925-065A-428890
Query Match          74.0%; Score 222; DB 6; Length 627;
Best Local Similarity 99.1%; Pred. No. 9.5e-50;
Matches 222; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGCATTGACAGGCTAAATGCTAAGTACTAGTACGATGAGACTTGAATATCTTCATTGGA 60
Db 224 GGCATTGACAGGCTAAATGCTAAGTACTAGTACGATGAGACTTGAATATCTTCATTGGA 165

Qy 61 AATAAAGCTTAATTTTCAATTAAGAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 120
Db 164 AATAAAGCTTAATTTTCAATTAAGAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 105

Qy 121 AAATGCTTAATTTTCAATTAAGAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 180
Db 104 AAATGCTTAATTTTCAATTAAGAGTAAACAGAGTGTGCAACATGCTGTTAAATAACCC 45

Qy 181 GACAACTTCAATCACTATAGTCTGTAGTAGTGCATCTGCAAGGATCCAGAGTAACC 224
Db 44 GACAACTTCAATCACTATAGTCTGTAGTAGTGCATCTGCAAGGATCCAGAGTAACC 1

RESULT 5
US-10-027-632-116164/c
; Sequence 116164; Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
```

```

; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 116165
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-116165

Query Match      74.0%; Score 222; DB 6; Length 627;
Best Local Similarity 99.1%; Pred. No. 9.5e-50;
Matches 222; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1  GGCATTGACAGGCTAAATGCTAAGTGAAGTACGATGAGAGCTTGAATATTTCTTCATTGA 60
224  GGCATTGACAGGCTAAATGCTAAGTGAAGTACGATGAGAGCTTGAATATTTCTTCATTGA 165

61  AATAAAGAGCTAATCTCCAAATTTGCTAACTAGGAGGAGTTTATCATGAAGAAATGTTT 120
164  AATAAAGAGCTAATCTCCAAATTTGCTAACTAGGAGGAGTTTATCATGAAGAAATGTTT 105

121  AATGCTTAAATTTCAATTAAGAGTAACAGAGTGTGCAACATGCTGTTAAATAACCC 180
104  AATGCTTAAATTTCAATTAAGAGTAACAGAGTGTGCAACATGCTGTTAAATAACCC 45

181  GACAACTTCAATCACTATAGCTGTAGTAGAGTGCAATCTGCAA 224
44  GACAACTTCAATCACTATAGCTGTAGTAGAGTGCAATCTGCAA 1

RESULT 6
US-10-027-632-116164/c
; Sequence 116164, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 116164
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-116165

Query Match      74.0%; Score 222; DB 6; Length 627;
Best Local Similarity 99.1%; Pred. No. 9.5e-50;
Matches 222; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1  GGCATTGACAGGCTAAATGCTAAGTGAAGTACGATGAGAGCTTGAATATTTCTTCATTGA 60
224  GGCATTGACAGGCTAAATGCTAAGTGAAGTACGATGAGAGCTTGAATATTTCTTCATTGA 165

61  AATAAAGAGCTAATCTCCAAATTTGCTAACTAGGAGGAGTTTATCATGAAGAAATGTTT 120
164  AATAAAGAGCTAATCTCCAAATTTGCTAACTAGGAGGAGTTTATCATGAAGAAATGTTT 105

121  AATGCTTAAATTTCAATTAAGAGTAACAGAGTGTGCAACATGCTGTTAAATAACCC 180
104  AATGCTTAAATTTCAATTAAGAGTAACAGAGTGTGCAACATGCTGTTAAATAACCC 45

181  GACAACTTCAATCACTATAGCTGTAGTAGAGTGCAATCTGCAA 224
44  GACAACTTCAATCACTATAGCTGTAGTAGAGTGCAATCTGCAA 1

RESULT 6
US-10-027-632-116164/c
; Sequence 116164, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 116164
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-116165
```

```

; ORGANISM: Human
; US-10-027-632-116164

Query Match      74.0%; Score 222; DB 7; Length 627;
Best Local Similarity 99.1%; Pred. No. 9.5e-50;
Matches 222; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1  GGCATTGACAGGCTAAATGCTAAGTGAAGTACGATGAGAGCTTGAATATTTCTTCATTGA 60
224  GGCATTGACAGGCTAAATGCTAAGTGAAGTACGATGAGAGCTTGAATATTTCTTCATTGA 165

61  AATAAAGAGCTAATCTCCAAATTTGCTAACTAGGAGGAGTTTATCATGAAGAAATGTTT 120
164  AATAAAGAGCTAATCTCCAAATTTGCTAACTAGGAGGAGTTTATCATGAAGAAATGTTT 105

121  AATGCTTAAATTTCAATTAAGAGTAACAGAGTGTGCAACATGCTGTTAAATAACCC 180
104  AATGCTTAAATTTCAATTAAGAGTAACAGAGTGTGCAACATGCTGTTAAATAACCC 45

181  GACAACTTCAATCACTATAGCTGTAGTAGAGTGCAATCTGCAA 224
44  GACAACTTCAATCACTATAGCTGTAGTAGAGTGCAATCTGCAA 1

RESULT 7
US-10-027-632-116165/c
; Sequence 116165, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 116165
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-116165

Query Match      74.0%; Score 222; DB 7; Length 627;
Best Local Similarity 99.1%; Pred. No. 9.5e-50;
Matches 222; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1  GGCATTGACAGGCTAAATGCTAAGTGAAGTACGATGAGAGCTTGAATATTTCTTCATTGA 60
224  GGCATTGACAGGCTAAATGCTAAGTGAAGTACGATGAGAGCTTGAATATTTCTTCATTGA 165

61  AATAAAGAGCTAATCTCCAAATTTGCTAACTAGGAGGAGTTTATCATGAAGAAATGTTT 120
164  AATAAAGAGCTAATCTCCAAATTTGCTAACTAGGAGGAGTTTATCATGAAGAAATGTTT 105

121  AATGCTTAAATTTCAATTAAGAGTAACAGAGTGTGCAACATGCTGTTAAATAACCC 180
104  AATGCTTAAATTTCAATTAAGAGTAACAGAGTGTGCAACATGCTGTTAAATAACCC 45
```

QY 181 GACAACTTCAATCACTAGCTAGTAGAGTGCATTCTGCA 224
DB |||||
44 GACAACTTCAATCACTAGCTAGTAGAGTGCATTCTGCA 1

RESULT 8

US-10-995-561-33431/c
; Sequence 33431, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33431
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-33431

Query Match 13.5%; Score 40.6; DB 10; Length 201;
Best Local Similarity 55.5%; Pred. No. 1;
Matches 76; Conservative 1; Mismatches 60; Indels 0; Gaps 0;

QY 18 TGCTAAGTACGATGACGACTTGAATATCTTCAATTTGAATAAAGCTAACTCC 77
DB |||||
188 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTAATAA 129
QY 78 CAAATTGTACCTAAGTGGAGTTTATCATGAAGAAATGTTAAATGCTTAATTTCAA 137
DB |||||
128 CAGAAGTTTAAAGTAAAGAGCAAAAYAAAGAGAAATGATGAACAAATATGTAA 69
QY 138 ATAAGAAGTAAACAGAG 154
DB |||||
68 CTAAGAAATAAAGAGAG 52

RESULT 9

US-10-995-561-37932/c
; Sequence 37932, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37932
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-37932

Query Match 13.5%; Score 40.6; DB 10; Length 201;
Best Local Similarity 55.5%; Pred. No. 1;
Matches 76; Conservative 1; Mismatches 60; Indels 0; Gaps 0;

QY 18 TGCTAAGTACGATGACGACTTGAATATCTTCAATTTGAATAAAGCTAACTCC 77
DB |||||
188 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTAATAA 129
QY 78 CAAATTGTACCTAAGTGGAGTTTATCATGAAGAAATGTTAAATGCTTAATTTCAA 137
DB |||||
128 CAGAAGTTTAAAGTAAAGAGCAAAAYAAAGAGAAATGATGAACAAATATGTAA 69

QY 138 ATAAGAAGTAAACAGAG 154
DB |||||
68 CTAAGAAATAAAGAGAG 52

RESULT 10

US-09-925-065A-349719
; Sequence 349719, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349719
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-349719

Query Match 13.5%; Score 40.6; DB 4; Length 588;
Best Local Similarity 55.5%; Pred. No. 1.6;
Matches 76; Conservative 1; Mismatches 60; Indels 0; Gaps 0;

QY 18 TGCTAAGTACGATGACGACTTGAATATCTTCAATTTGAATAAAGCTAACTCC 77
DB |||||
109 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTAATAA 168
QY 78 CAAATTGTACCTAAGTGGAGTTTATCATGAAGAAATGTTAAATGCTTAATTTCAA 137
DB |||||
169 CAGAAGTTTAAAGTAAAGAGCAAAAYAAAGAGAAATGATGAACAAATATGTAA 228
QY 138 ATAAGAAGTAAACAGAG 154
DB |||||
229 CTAAGAAATAAAGAGAG 245

RESULT 11

US-09-925-065A-349719
; Sequence 349719, Application US/09925065A
; Publication No. US2005028172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09


```

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13277
; LENGTH: 173995
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-995-561-13277

Query Match      13.5%; Score 40.6; DB 10; Length 173995;
Best Local Similarity 55.5%; Pred. No. 18;
Matches 76; Conservative 1; Mismatches 60; Indels 0; Gaps 0;

QY 18 TCGTAAGTACTACGATGAAGACTTGAAATATCTTCATTTGAAATAAACAGCTTAACCTCC 77
DB 28061 TACATAGTGAAGTCAAGAGAGAACTTTAAATGTTTACAGATCCCAAAATATTAATAATTTAA 28002

QY 78 CAAATTCTACTACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTCAA 137
DB 28001 CAGAAGTTATTAAGTAAAGAGAGCAAAAYAAAGAAAGAAATGAGTAAACAAATATGTAA 27942

QY 138 ATAGAAGTAAACAGAG 154
DB 27941 CTAAATAAAGAGAG 27925

RESULT 16
US-10-301-480-21248
; Sequence 21248, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21248
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapien
; US-10-301-480-21248

Query Match      13.2%; Score 39.6; DB 12; Length 483;
Best Local Similarity 51.8%; Pred. No. 2.8;
Matches 87; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

QY 74 CTCCAAATTTGCTACCTAAGAGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTT 133
DB 8 CTACTCCATTTGGATCTTAACCTTTGAGATACCAAGAAATAATATTTGAGCATATCCTTATAGG 67

QY 134 TCAAATAGAGTAACACAGAGTGTTCGAACATGCTGTTAAATACCCGACAACTTCAAT 193
DB 68 TTAGATACGGAGTTACATCTGACATATAATGAGTGATACATAAGAAAGTAGACTGAAAT 127

QY 194 CACTATAGCTGTAGTAGTGAGTGCAATCTTGCAAGGATCCAGAGTAACCA 241
DB 128 ACATACAGCATTTGAGTATAGAAGTTTCAAGGGTTCAAGGGTAGCAA 175

RESULT 17
US-10-301-480-634657
; Sequence 634657, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 634657
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapien
; US-10-301-480-634657

Query Match      13.2%; Score 39.6; DB 12; Length 483;
Best Local Similarity 51.8%; Pred. No. 2.8;
Matches 87; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

QY 74 CTCCAAATTTGCTACCTAAGAGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTT 133
DB 8 CTACTCCATTTGGATCTTAACCTTTGAGATACCAAGAAATAATATTTGAGCATATCCTTATAGG 67

QY 134 TCAAATAGAGTAACACAGAGTGTTCGAACATGCTGTTAAATACCCGACAACTTCAAT 193
DB 68 TTAGATACGGAGTTACATCTGACATATAATGAGTGATACATAAGAAAGTAGACTGAAAT 127

QY 194 CACTATAGCTGTAGTAGTGAGTGCAATCTTGCAAGGATCCAGAGTAACCA 241
DB 128 ACATACAGCATTTGAGTATAGAAGTTTCAAGGGTTCAAGGGTAGCAA 175

RESULT 18
US-10-027-632-93552
; Sequence 93552, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93552
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-93552
```

Query Match 13.2%; Score 39.6; DB 6; Length 502;
Best Local Similarity 51.8%; Pred. No. 2.9;
Matches 87; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

QY 74 CTCCTCAATGTTACCTAAGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTT 133
DB 7 CTACTCCATTGGATCTAACTTTTGATATACCAAGAAATATATTTGAGCATATCCTTATAGG 66

QY 134 TCAATAAAGAAAGTAACAGAGTGTTCACATGCTGTTAAATTAACCCGACAAACTTCAAT 193
DB 67 TTAGATACGGAGTTACATCTGACATATAATGAGTGTAACAATAAGAAAGTAGACTGAAAT 126

QY 194 CACTATAGCTGTAGTAGAGTGCTTCTGCAAGGATCCAGAGTAACCA 241
DB 127 ACATACAGCATTTGAGTATAGAAGTTTTCAGGGGTTTCAAGGGTAGCAA 174

RESULT 19

US-10-027-632-305099
; Sequence 305099, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 305099
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-305099

Query Match 13.2%; Score 39.6; DB 6; Length 502;
Best Local Similarity 51.8%; Pred. No. 2.9;
Matches 87; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

QY 74 CTCCTCAATGTTACCTAAGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTT 133
DB 7 CTACTCCATTGGATCTAACTTTTGATATACCAAGAAATATATTTGAGCATATCCTTATAGG 66

QY 134 TCAATAAAGAAAGTAACAGAGTGTTCACATGCTGTTAAATTAACCCGACAAACTTCAAT 193
DB 67 TTAGATACGGAGTTACATCTGACATATAATGAGTGTAACAATAAGAAAGTAGACTGAAAT 126

QY 194 CACTATAGCTGTAGTAGAGTGCTTCTGCAAGGATCCAGAGTAACCA 241
DB 127 ACATACAGCATTTGAGTATAGAAGTTTTCAGGGGTTTCAAGGGTAGCAA 174

RESULT 20

US-10-027-632-93552
; Sequence 93552, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93552
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-93552

Query Match 13.2%; Score 39.6; DB 7; Length 502;
Best Local Similarity 51.8%; Pred. No. 2.9;
Matches 87; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

QY 74 CTCCTCAATGTTACCTAAGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTT 133
DB 7 CTACTCCATTGGATCTAACTTTTGATATACCAAGAAATATATTTGAGCATATCCTTATAGG 66

QY 134 TCAATAAAGAAAGTAACAGAGTGTTCACATGCTGTTAAATTAACCCGACAAACTTCAAT 193
DB 67 TTAGATACGGAGTTACATCTGACATATAATGAGTGTAACAATAAGAAAGTAGACTGAAAT 126

QY 194 CACTATAGCTGTAGTAGAGTGCTTCTGCAAGGATCCAGAGTAACCA 241
DB 127 ACATACAGCATTTGAGTATAGAAGTTTTCAGGGGTTTCAAGGGTAGCAA 174

RESULT 21

US-10-027-632-305099
; Sequence 305099, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 305099
; LENGTH: 502

```
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-305099

Query Match      13.2%; Score 39.6; DB 7; Length 502;
Best Local Similarity 51.8%; Pred. No. 2.9;
Matches 87; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

QY 74 CTCACAAATTTGACTTAAGTGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTT 133
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7 CTACTCCATTTGGATCTTAATTTGAGATACCAAGAAATAATATTGAGCATATCTCTTATAGG 66

QY 134 TCAATAAGAGACTACCAAGAGTTTGCACATGCTTTAATAAATACCCGACAACTTCAAT 193
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 67 TTAGATCGGAGTTTACATCTGCATATATATGAGTGTAACAATAAGAAAGTAGACTGNAAT 126

QY 194 CACTATAGCTGTAGTAGAGTGATCTCTGCAAGGATCCGAGAGTAACCA 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 127 ACATACAGCATTTGAGTAGTAGAAGTTTTCAGGGGTTCAAGGGTAGCAA 174
```

```
RESULT 22
US-09-925-065A-349720
; Sequence 349720, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349720
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-349720
```

```
Query Match      13.1%; Score 39.4; DB 4; Length 588;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTGACTACGATGAAGACTTGAATAATTTCTTCAATTTGAATAACAGCTAACTCC 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 109 TACATAGTGAAGTGAAGAGAACTTTAATGTTTTTACAGATCCCAATATTAAAAATTTAA 168

QY 78 CAAATTTGACTTAAGTGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAA 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 169 CAGAAAGTTATTAAAGTAAAGAGCAACAAAGAGAAATGATGTAACAAATATGTAA 228

QY 138 ATAAGAAGTACACAGAG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 229 CTAAAAAATAAGAGAG 245
```

```
RESULT 23
US-09-925-065A-349722
; Sequence 349722, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```

```
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349722
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-349722
```

```
Query Match      13.1%; Score 39.4; DB 4; Length 588;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTGACTACGATGAAGACTTGAATAATTTCTTCAATTTGAATAACAGCTAACTCC 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 109 TACATAGTGAAGTGAAGAGAACTTTAATGTTTTTACAGATCCCAATATTAAAAATTTAA 168

QY 78 CAAATTTGACTTAAGTGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAA 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 169 CAGAAAGTTATTAAAGTAAAGAGCAACAAAGAGAAATGATGTAACAAATATGTAA 228

QY 138 ATAAGAAGTACACAGAG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 229 CTAAAAAATAAGAGAG 245
```

```
RESULT 24
US-09-925-065A-349720
; Sequence 349720, Application US/09925065A
; Publication No. US2005028172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349720
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-349720
```

```
Query Match      13.1%; Score 39.4; DB 5; Length 588;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
```

| | | | |
|----|-----|----------------------------------------------------------------|-----|
| Qy | 18 | TGCTAAGTCACTACGATGAGAGCTTGGAAATATCTTCATTTTGAAGTAAACAGCTAACTCC | 77 |
| | | | |
| Db | 109 | TACATAGTGAAGTGAAGAGAGAACTTTAAATGTTTTACAGATCCAAATATTAAAAATTTAA | 168 |
| Qy | 78 | CAAAATTGTTACCTAAGTGGGGAGTTTATCATGAGAGAAATGTTTAAATGCTTAATTTTCAA | 137 |
| | | | |
| Db | 169 | CAGAAGTTATTAAAGTTAAAGAGCAACAAAGAGAAATGATGATAAAAACAATAATGTAA | 228 |
| Qy | 138 | ATGAAGAGTTAACCGAG | 154 |
| | | | |
| Db | 229 | CTAAAAAATAAAGAGAG | 245 |

```

RESULT 25
US-09-925-065A-349722
; Sequence 349722, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827,135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349722
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-349722

```

```

RESULT 26
US-10-301-480-421780
: Sequence 421780, Application US/10301480
: Publication No. US20060057564A1
: GENERAL INFORMATION:
: APPLICANT: Wang, David G.
: TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
: TITLE OF INVENTION: in the Human Genome
: FILE REFERENCE: 108827.137
: CURRENT APPLICATION NUMBER: US/10/301,480
: CURRENT FILING DATE: 2002-11-21
: PRIOR APPLICATION NUMBER: US 10/215,598
: PRIOR FILING DATE: 2002-08-09
: PRIOR APPLICATION NUMBER: US 60/311,695

```

```

; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows
; SEQ ID NO 421780
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-421780

```

```

RESULT 27
US-10-301-480-421782
; Sequence 421782, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 421782
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-421782

```

RESULT 28
US-10-301-480-1035189
; Sequence 1035189, Application US/10301480
; Publication NO. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: wang, David G.
; TITLE OF INVENTION: Identification and Ma


```
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1035189
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1035189

Query Match      13.1%; Score 39.4; DB 12; Length 599;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTGACTACGATGAAGACTTGAATAATCTTCAATTTGAAATAAACAGCTAACTCC 77
DB 109 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTTAAATTTAA 168
QY 78 CAAATTGTACCTAACTAGGAGCTTTATCATGAAGAAATGTTTAAATGCTTTAAATTTCAA 137
DB 169 CAGAAGTTATTAAAGTAAAGAGCAAAACAAAGAGAAATGATGATGATGATGATGATGAT 228
QY 138 ATAAGAAGTAAACACAG 154
DB 229 CTAAATAAATAAGAGAG 245

RESULT 29
US-10-301-480-1035191
; Sequence 1035191, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1035191
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1035191

Query Match      13.1%; Score 39.4; DB 12; Length 599;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTGACTACGATGAAGACTTGAATAATCTTCAATTTGAAATAAACAGCTAACTCC 77
DB 109 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTTAAATTTAA 168
QY 78 CAAATTGTACCTAACTAGGAGCTTTATCATGAAGAAATGTTTAAATGCTTTAAATTTCAA 137
DB 169 CAGAAGTTATTAAAGTAAAGAGCAAAACAAAGAGAAATGATGATGATGATGATGATGAT 228
QY 138 ATAAGAAGTAAACACAG 154
DB 229 CTAAATAAATAAGAGAG 245
```

```
RESULT 30
US-09-925-065A-349721
; Sequence 349721, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349721
; LENGTH: 602
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-349721

Query Match      13.1%; Score 39.4; DB 4; Length 602;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTGACTACGATGAAGACTTGAATAATCTTCAATTTGAAATAAACAGCTAACTCC 77
DB 58 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTACAGATCCAAATATTTAAATTTAA 117
QY 78 CAAATTGTACCTAACTAGGAGCTTTATCATGAAGAAATGTTTAAATGCTTTAAATTTCAA 137
DB 118 CAGAAGTTATTAAAGTAAAGAGCAAAACAAAGAGAAATGATGATGATGATGATGATGAT 177
QY 138 ATAAGAAGTAAACACAG 154
DB 178 CTAAATAAATAAGAGAG 194

RESULT 31
US-09-925-065A-349721
; Sequence 349721, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349721
; LENGTH: 602
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
US-09-925-065A-349721
Query Match      13.1%; Score 39.4; DB 5; Length 602;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTACTACGATGAAGACTTGAATATCTTCATTGTAATAAACAAGCTAACTCC 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 58 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTTACAGATCCAAATATTTAAAAATTTAA 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 78 CAAATTGTACTTAAGTGAAGAGGAGTTTATCATGAAGAATGTTTAAATGCTTTAAATTTCAA 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 CAGAGTTATTAAAGTAAAGAGCGCAACAAGAGAAGAAATGAGTAAACAATAATGTAA 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 138 ATAAGAAGTAACCCAGAG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 CTAAAAATAAAGAGAG 194
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 32
US-10-301-480-421781
; Sequence 421781, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-08-09
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 421781
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-421781

Query Match      13.1%; Score 39.4; DB 12; Length 608;
Best Local Similarity 55.5%; Pred. No. 3.5;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTACTACGATGAAGACTTGAATATCTTCATTGTAATAAACAAGCTAACTCC 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 58 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTTACAGATCCAAATATTTAAAAATTTAA 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 78 CAAATTGTACTTAAGTGAAGAGGAGTTTATCATGAAGAATGTTTAAATGCTTTAAATTTCAA 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 CAGAGTTATTAAAGTAAAGAGCGCAACAAGAGAAGAAATGAGTAAACAATAATGTAA 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 138 ATAAGAAGTAACCCAGAG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 CTAAAAATAAAGAGAG 194
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 33
US-10-301-480-1035190
; Sequence 1035190, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-08-09
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1035190
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1035190

Query Match      13.1%; Score 39.4; DB 7; Length 60461;
Best Local Similarity 55.5%; Pred. No. 25;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 18 TGCTAAGTACTACGATGAAGACTTGAATATCTTCATTGTAATAAACAAGCTAACTCC 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 27844 TACATAGTGAAGTGAAGAGAACTTTAAATGTTTTACAGATCCAAATATTTAAAAATTTAA 27785
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 78 CAAATTGTACTTAAGTGAAGAGGAGTTTATCATGAAGAATGTTTAAATGCTTTAAATTTCAA 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 27784 CAGAGTTATTAAAGTAAAGAGCGCAACAAGAGAAGAAATGAGTAAACAATAATGTAA 27725
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 138 ATAAGAAGTAACCCAGAG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 27724 CTAAAAATAAAGAGAG 27708
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 35
US-10-087-192-331
; Sequence 331, Application US/10087192
; Publication No. US20020182586A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 331
; LENGTH: 295096
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(295096)
; OTHER INFORMATION: n = A,T,C or G
; US-10-087-192-331.

Query Match      12.9%; Score 38.8; DB 6; Length 295096;
Best Local Similarity 48.9%; Pred. No. 71;
Matches 132; Conservative 0; Mismatches 137; Indels 1; Gaps 1;

1  GGCATTGACAGGCTAAATGCTAAGTGACTACGATGACGACTTGAATAATTTCTTCATTGA 60
175844  GCCATTGCTGAAGACTCTGCTGAGGCTAGATGAAGAGATTCATTAATTCGTTGA 175903
61  AATAAACAGCTPACTCCAAATTTGACCTAACTAGGGGAGTTTATCATGAAGAATGTTT 120
175904  AAAAGAAAGCCTCAAAACAGCTCTGTATAAATTTTGTGTTTATTCAGTTCACTCTTA 175963
121  AATGCTTAATTTTCAATATAGAGTAACAGAGTGTGCAACATGCTGT-TAAATACC 179
175964  TGAAGGAATGTTTAAATAAAGAGGATGAGCTGAGAAAGGAAAAATACAAAATATATGGCT 176023
180  CGACAAACTTCAATCACTATAGCTGTAGTAGAGTGCTTCTGCAAGGATCCAGAGTAAC 239
176024  CAAGTATTAAGGGGCATCAAGAAGTAGAATGGAGCTGAATCCTATGTTCCAGGAGTTAA 176083
240  CAGTATTTTGGAAATGCAATGTTGAACCGA 269
176084  CAAAATAGGAGTGTTGTAATCTTGAGGCAA 176113

RESULT 36
US-10-357-930-21832
; Sequence 21832, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22160
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1_2_3_4_5_6_2220_2221_2222_2223_2224
; OTHER INFORMATION: n = A,T,C or G
; US-10-357-930-22160

```

```

; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21832
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1_2_3_4_5_6_2220_2221_2222_2223_2224
; OTHER INFORMATION: n = A,T,C or G
; US-10-357-930-21832

Query Match      12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

Qy  41  TTGAATATTTCTTCATTGAAATAAACAGCCTAACTCCCAATTTGTACCTAAGGGAG 100
      |||||
Db  2035 TTGAATATTTTAAAGCTTAAATTTCTGCTGTTTGCARATTTGATATAACATTATATC 2094
101  TTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAAATGAAGTAACAGAGTGTG 159
      |||||
Db  2095 TATATTATTAACCTGAAGTATTTTTCCTGCTGATTATGAATATAAAGTAACACAGGGTTA 2154
160  CAACATGCTGTTAATAAACCGACAAACTTCAATCACTATA 200
      |||||
Db  2155 CAGTTGTGCAAAACAAAACATACCTAAATTAACAATA 2195

RESULT 37
US-10-357-930-22160
; Sequence 22160, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22160
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1_2_3_4_5_6_2220_2221_2222_2223_2224
; OTHER INFORMATION: n = A,T,C or G
; US-10-357-930-22160

```

Query Match 12.9% Score 38.6; DB 9; Length 2224;
 Best Local Similarity 56.5%; Pred. No. 10;
 Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
 QY 41 TTGAAATATTCTTCATTTGAAATAAACAGCTAACTCCCAAAATGTTACTTAACCTAGGGGAG 100
 DB 2035 TTGAAATATTCTTCATTTGAAATAAACAGCTAACTCCCAAAATGTTACTTAACCTATATC 2094
 QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAAAATGAAGTAACAGAGTGTG 159
 DB 2095 TATATTATTAAGTGAAGTATTTTTCCTGATTTATGAAATATAAGTAACACAGGGTTA 2154
 QY 160 CAACATGCTGTTAAATAACCGCAAACTTCAATCACTATA 200
 DB 2155 CAGTTGTGCAAAACAAACACATACCTAAATTAACAATA 2195

RESULT 38
 US-10-357-930-22166
 ; Sequence 22166, Application US/10357930
 ; Publication No. US20040259086A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Schlegel, Robert
 ; APPLICANT: Endege, Wilson
 ; APPLICANT: Monahan, John
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
 ; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
 ; TITLE OF INVENTION: HUMAN PROSTATE CANCER
 ; FILE REFERENCE: MRI-007BCN
 ; CURRENT APPLICATION NUMBER: US/10/357,930
 ; CURRENT FILING DATE: 2003-02-04
 ; PRIOR APPLICATION NUMBER: 09/785,276
 ; PRIOR FILING DATE: 2003-02-16
 ; PRIOR APPLICATION NUMBER: 60/183,319
 ; PRIOR FILING DATE: 2000-02-17
 ; PRIOR APPLICATION NUMBER: 60/189,862
 ; PRIOR FILING DATE: 2000-03-16
 ; PRIOR APPLICATION NUMBER: 60/207,454
 ; PRIOR FILING DATE: 2000-05-25
 ; PRIOR APPLICATION NUMBER: 60/211,314
 ; PRIOR FILING DATE: 2000-06-09
 ; PRIOR APPLICATION NUMBER: 60/219,007
 ; PRIOR FILING DATE: 2000-07-18
 ; PRIOR APPLICATION NUMBER: 60/255,281
 ; PRIOR FILING DATE: 2000-12-13
 ; NUMBER OF SEQ ID NOS: 62232
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 22166
 ; LENGTH: 2224
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc feature
 ; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-10-357-930-22166

Query Match 12.9% Score 38.6; DB 9; Length 2224;
 Best Local Similarity 56.5%; Pred. No. 10;
 Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
 QY 41 TTGAAATATTCTTCATTTGAAATAAACAGCTAACTCCCAAAATGTTACTTAACCTAGGGGAG 100
 DB 2035 TTGAAATATTCTTCATTTGAAATAAACAGCTAACTCCCAAAATGTTACTTAACCTATATC 2094
 QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAAAATGAAGTAACAGAGTGTG 159
 DB 2095 TATATTATTAAGTGAAGTATTTTTCCTGATTTATGAAATATAAGTAACACAGGGTTA 2154
 QY 160 CAACATGCTGTTAAATAACCGCAAACTTCAATCACTATA 200
 DB 2155 CAGTTGTGCAAAACAAACACATACCTAAATTAACAATA 2195

RESULT 39
 US-10-357-930-23173
 ; Sequence 23173, Application US/10357930
 ; Publication No. US20040259086A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Schlegel, Robert
 ; APPLICANT: Endege, Wilson
 ; APPLICANT: Monahan, John
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
 ; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
 ; TITLE OF INVENTION: HUMAN PROSTATE CANCER
 ; FILE REFERENCE: MRI-007BCN
 ; CURRENT APPLICATION NUMBER: US/10/357,930
 ; CURRENT FILING DATE: 2003-02-04
 ; PRIOR APPLICATION NUMBER: 09/785,276
 ; PRIOR FILING DATE: 2003-02-16
 ; PRIOR APPLICATION NUMBER: 60/183,319
 ; PRIOR FILING DATE: 2000-02-17
 ; PRIOR APPLICATION NUMBER: 60/189,862
 ; PRIOR FILING DATE: 2000-03-16
 ; PRIOR APPLICATION NUMBER: 60/207,454
 ; PRIOR FILING DATE: 2000-05-25
 ; PRIOR APPLICATION NUMBER: 60/211,314
 ; PRIOR FILING DATE: 2000-06-09
 ; PRIOR APPLICATION NUMBER: 60/219,007
 ; PRIOR FILING DATE: 2000-07-18
 ; PRIOR APPLICATION NUMBER: 60/255,281
 ; PRIOR FILING DATE: 2000-12-13
 ; NUMBER OF SEQ ID NOS: 62232
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 23173
 ; LENGTH: 2224
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc feature
 ; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-10-357-930-23173

Query Match 12.9% Score 38.6; DB 9; Length 2224;
 Best Local Similarity 56.5%; Pred. No. 10;
 Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
 QY 41 TTGAAATATTCTTCATTTGAAATAAACAGCTAACTCCCAAAATGTTACTTAACCTAGGGGAG 100
 DB 2035 TTGAAATATTCTTCATTTGAAATAAACAGCTAACTCCCAAAATGTTACTTAACCTATATC 2094
 QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAAAATGAAGTAACAGAGTGTG 159
 DB 2095 TATATTATTAAGTGAAGTATTTTTCCTGATTTATGAAATATAAGTAACACAGGGTTA 2154
 QY 160 CAACATGCTGTTAAATAACCGCAAACTTCAATCACTATA 200
 DB 2155 CAGTTGTGCAAAACAAACACATACCTAAATTAACAATA 2195

RESULT 40
 US-10-357-930-24502
 ; Sequence 24502, Application US/10357930
 ; Publication No. US20040259086A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Schlegel, Robert
 ; APPLICANT: Endege, Wilson
 ; APPLICANT: Monahan, John
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
 ; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
 ; TITLE OF INVENTION: HUMAN PROSTATE CANCER
 ; FILE REFERENCE: MRI-007BCN
 ; CURRENT APPLICATION NUMBER: US/10/357,930
 ; CURRENT FILING DATE: 2003-02-04
 ; PRIOR APPLICATION NUMBER: 09/785,276

;; PRIOR FILING DATE: 2003-02-16
;; PRIOR APPLICATION NUMBER: 60/183,319
;; PRIOR FILING DATE: 2000-02-17
;; PRIOR APPLICATION NUMBER: 60/189,862
;; PRIOR FILING DATE: 2000-03-16
;; PRIOR APPLICATION NUMBER: 60/207,454
;; PRIOR FILING DATE: 2000-05-25
;; PRIOR APPLICATION NUMBER: 60/211,314
;; PRIOR FILING DATE: 2000-06-09
;; PRIOR APPLICATION NUMBER: 60/219,007
;; PRIOR FILING DATE: 2000-07-18
;; PRIOR APPLICATION NUMBER: 60/255,281
;; PRIOR FILING DATE: 2000-12-13
;; NUMBER OF SEQ ID NOS: 62232
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 24502
;; LENGTH: 2224
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; NAME/KEY: misc_feature
;; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
;; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24502

Query Match 12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
QY 41 TTGAATATTTCTTCATTTGAAATAAAGCTTAAATTCGCTGGTTGCAAAATTTGTAATAACATTATATC 2094
DB 2035 TTGAATATTTTAAAGCTTAAATTCGCTGGTTGCAAAATTTGTAATAACATTATATC 2094
QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAAAATGAAGTAACAGAGTGTG 159
DB 2095 TATATTATTAAGTAAATTTTTCCTGATTAATGAATATGAAGTAACAGAGGTTA 2154
QY 160 CAACATGCTGTTAAATAACCCGACAACTTCAATCACTATA 200
DB 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2195

RESULT 41
US-10-357-930-24579
;; Sequence 24579, Application US/10357930
;; Publication No. US20040259086A1
;; GENERAL INFORMATION:
;; APPLICANT: Schlegel, Robert
;; APPLICANT: Endege, Wilson
;; APPLICANT: Monahan, John
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;; FILE REFERENCE: MRI-007BCN
;; CURRENT APPLICATION NUMBER: US/10/357,930
;; CURRENT FILING DATE: 2003-02-04
;; PRIOR FILING DATE: 2003-02-16
;; PRIOR APPLICATION NUMBER: 60/183,319
;; PRIOR FILING DATE: 2000-02-17
;; PRIOR APPLICATION NUMBER: 60/189,862
;; PRIOR FILING DATE: 2000-03-16
;; PRIOR APPLICATION NUMBER: 60/207,454
;; PRIOR FILING DATE: 2000-05-25
;; PRIOR APPLICATION NUMBER: 60/211,314
;; PRIOR FILING DATE: 2000-06-09
;; PRIOR APPLICATION NUMBER: 60/219,007
;; PRIOR FILING DATE: 2000-07-18
;; PRIOR APPLICATION NUMBER: 60/255,281
;; NUMBER OF SEQ ID NOS: 62232
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 24579

;; LENGTH: 2224
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; NAME/KEY: misc_feature
;; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
;; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24579

Query Match 12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
QY 41 TTGAATATTTCTTCATTTGAAATAAAGCTTAAATTCGCTGGTTGCAAAATTTGTAATAACATTATATC 100
DB 2035 TTGAATATTTTAAAGCTTAAATTCGCTGGTTGCAAAATTTGTAATAACATTATATC 2094
QY 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAAAATGAAGTAACAGAGTGTG 159
DB 2095 TATATTATTAAGTAAATTTTTCCTGATTAATGAATATGAAGTAACAGAGGTTA 2154
QY 160 CAACATGCTGTTAAATAACCCGACAACTTCAATCACTATA 200
DB 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2195

RESULT 42
US-10-357-930-24590
;; Sequence 24590, Application US/10357930
;; Publication No. US20040259086A1
;; GENERAL INFORMATION:
;; APPLICANT: Schlegel, Robert
;; APPLICANT: Endege, Wilson
;; APPLICANT: Monahan, John
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;; FILE REFERENCE: MRI-007BCN
;; CURRENT APPLICATION NUMBER: US/10/357,930
;; CURRENT FILING DATE: 2003-02-04
;; PRIOR FILING DATE: 2003-02-16
;; PRIOR APPLICATION NUMBER: 60/183,319
;; PRIOR FILING DATE: 2000-02-17
;; PRIOR APPLICATION NUMBER: 60/189,862
;; PRIOR FILING DATE: 2000-03-16
;; PRIOR APPLICATION NUMBER: 60/207,454
;; PRIOR FILING DATE: 2000-05-25
;; PRIOR APPLICATION NUMBER: 60/211,314
;; PRIOR FILING DATE: 2000-06-09
;; PRIOR APPLICATION NUMBER: 60/219,007
;; PRIOR FILING DATE: 2000-07-18
;; PRIOR APPLICATION NUMBER: 60/255,281
;; NUMBER OF SEQ ID NOS: 62232
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 24590
;; LENGTH: 2224
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; NAME/KEY: misc_feature
;; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
;; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24590

Query Match 12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
QY 41 TTGAATATTTCTTCATTTGAAATAAAGCTTAAATTCGCTGGTTGCAAAATTTGTAATAACATTATATC 100
DB 2035 TTGAATATTTTAAAGCTTAAATTCGCTGGTTGCAAAATTTGTAATAACATTATATC 2094

```
Qy 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAAGTAAGTAACACAGAGTGTG 159
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2095 TATATTATTAACCTGAAGTATTTTGGCTGATTATGAATATAAAGTAACACAGGGTTA 2154

Qy 160 CAACATGCTGTTAAATACCGCAAACTTCAATCACTATA 200
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2195
```

```
RESULT 43
US-10-357-930-25046
; Sequence 25046, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25046
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-25046
```

```
Query Match 12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

Qy 41 TTGAATATTTCTTCATTTGAAATAAACAGCTAACTCCCAAAATGTACTTAACCTAGGGAG 100
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2035 TTGAATATTTTAAAGCTTAAATCTGCTGGTTTGCAAAATGTATATAACATTATATC 2094

Qy 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAAGTAAGTAACACAGAGTGTG 159
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2095 TATATTATTAACCTGAAGTATTTTGGCTGATTATGAATATAAAGTAACACAGGGTTA 2154

Qy 160 CAACATGCTGTTAAATACCGCAAACTTCAATCACTATA 200
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2195
```

```
RESULT 44
US-10-357-930-27679
; Sequence 27679, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
```

```
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27679
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-27679

Query Match 12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

Qy 41 TTGAATATTTCTTCATTTGAAATAAACAGCTAACTCCCAAAATGTACTTAACCTAGGGAG 100
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2035 TTGAATATTTTAAAGCTTAAATCTGCTGGTTTGCAAAATGTATATAACATTATATC 2094

Qy 101 TTTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAAGTAAGTAACACAGAGTGTG 159
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2095 TATATTATTAACCTGAAGTATTTTGGCTGATTATGAATATAAAGTAACACAGGGTTA 2154

Qy 160 CAACATGCTGTTAAATACCGCAAACTTCAATCACTATA 200
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2195
```

```
RESULT 45
US-10-357-930-28021
; Sequence 28021, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
```

PRIOR APPLICATION NUMBER: 60/211,314
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/219,007
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/255,281
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 62232
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28021
LENGTH: 2224
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
OTHER INFORMATION: n = A,T,C or G
US-10-357-930-28021

Query Match 12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
41 TTGAATATTTCTTCATTTGAATAAACAAGCTAACTCCCAAAATTTGACCTAACTAGGGGAG 100
|||||
2035 TTGAATATTTTAAAGCTTAAATCTGCTGTTGCAAAATTTGTAATATGATATAACATTATATC 2094
|||||
101 TTTATCATGAA-GAAATGTTTTAAATGCTTAATTTTCAAAATGAAGTAACAGAGTGTG 159
|||||
2095 TATATTATTAAGTCTGAAGTATTTTGGCTGATTATGAATATAAAGTAACACAGGGTTA 2154
|||||
160 CAACATGCTGTTAAATAACCGACAACTTCAATCACTATA 200
|||
2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2195
|||

RESULT 46

US-10-357-930-28027
Sequence 28027, Application US/10357930
Publication No. US20040259086A1
GENERAL INFORMATION:
APPLICANT: Schlengel, Robert
APPLICANT: Endege, Wilson
APPLICANT: Monahan, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
TITLE OF INVENTION: HUMAN PROSTATE CANCER
FILE REFERENCE: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211,314
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/219,007
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/255,281
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 62232
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28027
LENGTH: 2224
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
OTHER INFORMATION: n = A,T,C or G

US-10-357-930-28027

Query Match 12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
41 TTGAATATTTCTTCATTTGAATAAACAAGCTAACTCCCAAAATTTGACCTAACTAGGGGAG 100
|||||
2035 TTGAATATTTTAAAGCTTAAATCTGCTGTTGCAAAATTTGTAATATGATATAACATTATATC 2094
|||||
101 TTTATCATGAA-GAAATGTTTTAAATGCTTAATTTTCAAAATGAAGTAACAGAGTGTG 159
|||||
2095 TATATTATTAAGTCTGAAGTATTTTGGCTGATTATGAATATAAAGTAACACAGGGTTA 2154
|||||
160 CAACATGCTGTTAAATAACCGACAACTTCAATCACTATA 200
|||
2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2195
|||

RESULT 47

US-10-357-930-29041
Sequence 29041, Application US/10357930
Publication No. US20040259086A1
GENERAL INFORMATION:
APPLICANT: Schlengel, Robert
APPLICANT: Endege, Wilson
APPLICANT: Monahan, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
TITLE OF INVENTION: HUMAN PROSTATE CANCER
FILE REFERENCE: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211,314
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/219,007
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/255,281
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 62232
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 29041
LENGTH: 2224
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1, 2, 3, 4, 5, 6, 2220, 2221, 2222, 2223, 2224
OTHER INFORMATION: n = A,T,C or G
US-10-357-930-29041

Query Match 12.9%; Score 38.6; DB 9; Length 2224;
Best Local Similarity 56.5%; Pred. No. 10;
Matches 91; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
41 TTGAATATTTCTTCATTTGAATAAACAAGCTAACTCCCAAAATTTGACCTAACTAGGGGAG 100
|||||
2035 TTGAATATTTTAAAGCTTAAATCTGCTGTTGCAAAATTTGTAATATGATATAACATTATATC 2094
|||||
101 TTTATCATGAA-GAAATGTTTTAAATGCTTAATTTTCAAAATGAAGTAACAGAGTGTG 159
|||||
2095 TATATTATTAAGTCTGAAGTATTTTGGCTGATTATGAATATAAAGTAACACAGGGTTA 2154
|||||
160 CAACATGCTGTTAAATAACCGACAACTTCAATCACTATA 200
|||
2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2195
|||

```
Db 2155 CAGTTGTGCAAAACAAACAAACATACCTAAATTACAATA 2195

RESULT 48
US-10-027-632-86622/c
; Sequence 86622, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86622
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(552)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-86623

Query Match 12.7%; Score 38.2; DB 6; Length 552;
Best Local Similarity 56.0%; Pred. No. 7.2;
Matches 70; Conservative 1; Mismatches 54; Indels 0; Gaps 0;

Qy 108 TGAAGAAATGTTAAATGCTTAATTTCAATAAAGAGTAACAGAGTGTTCACATGC 167
Db 282 TAAATAATGTTAAAGGGTAATTTACAAAAAATGAATAATAAAGTGGACATGAAGT 223
Qy 168 TGTAAATAACCCGACAACTTCAATCACTATAGCTGTAGTAGAGTGCATTCTCAAGGA 227
Db 222 GATTAAATAACAGAAATCCAGAGCTATCATTATTGAYGTATGATATCTTCAATTTTAAAAA 163
Qy 228 TCCCA 232
Db 162 GGACA 158

RESULT 50
US-10-027-632-316419/c
; Sequence 316419, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 316419
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(552)
; OTHER INFORMATION: n = A,T,C or G

Query Match 12.7%; Score 38.2; DB 6; Length 552;
Best Local Similarity 56.0%; Pred. No. 7.2;
Matches 70; Conservative 1; Mismatches 54; Indels 0; Gaps 0;

Qy 108 TGAAGAAATGTTAAATGCTTAATTTCAATAAAGAGTAACAGAGTGTTCACATGC 167
Db 282 TAAATAATGTTAAAGGGTAATTTACAAAAAATGAATAATAAAGTGGACATGAAGT 223
Qy 168 TGTAAATAACCCGACAACTTCAATCACTATAGCTGTAGTAGAGTGCATTCTCAAGGA 227
Db 222 GATTAAATAACAGAAATCCAGAGCTATCATTATTGAYGTATGATATCTTCAATTTTAAAAA 163
Qy 228 TCCCA 232
Db 162 GGACA 158

RESULT 49
US-10-027-632-86623/c
; Sequence 86623, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
```


This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2006, 22:44:54 ; Search time 118 Seconds
(without alignments)
3544.714 Million cell updates/sec

Title: SEQ1-96535C

Perfect score: 300

Sequence: 1 ggcattgacaggttaatgc.....atctgctaggtaggaagtat 300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 866355 seqs, 697127050 residues

Total number of hits satisfying chosen parameters: 1772710

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA New.*

- 1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 9: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | Query Match | Length | ID | Description |
|--------|------|-------|-------------|--------|-----------------------|-------------------|
| 1 | 38.6 | 12.9 | 863 | 8 | US-11-266-748A-249055 | Sequence 249055, |
| 2 | 37.8 | 12.6 | 1000 | 8 | US-11-266-748A-201627 | Sequence 201627, |
| 3 | 37.8 | 12.6 | 1000 | 8 | US-11-266-748A-281554 | Sequence 281554, |
| 4 | 37.8 | 12.6 | 1000 | 8 | US-11-266-748A-308194 | Sequence 308194, |
| 5 | 37.8 | 12.6 | 1000 | 8 | US-11-266-748A-390539 | Sequence 390539, |
| 6 | 37.8 | 12.6 | 1000 | 8 | US-11-266-748A-481257 | Sequence 481257, |
| 7 | 37 | 12.3 | 6857 | 8 | US-11-266-748A-23772 | Sequence 23772, A |
| 8 | 36.8 | 12.3 | 692 | 8 | US-11-266-748A-16908 | Sequence 16908, A |
| 9 | 36.4 | 12.1 | 415 | 8 | US-11-266-748A-360446 | Sequence 360446, |
| 10 | 36.4 | 12.1 | 415 | 8 | US-11-266-748A-443825 | Sequence 443825, |
| 11 | 35.6 | 11.9 | 578 | 8 | US-11-266-748A-362773 | Sequence 362773, |
| 12 | 35.6 | 11.9 | 578 | 8 | US-11-266-748A-446152 | Sequence 446152, |
| 13 | 35.6 | 11.9 | 1000 | 8 | US-11-266-748A-403961 | Sequence 403961, |
| 14 | 35.6 | 11.9 | 1000 | 8 | US-11-266-748A-475007 | Sequence 475007, |
| 15 | 35.6 | 11.9 | 2205 | 8 | US-11-266-748A-31158 | Sequence 31158, A |
| 16 | 35.6 | 11.9 | 130263 | 8 | US-11-266-748A-58551 | Sequence 58551, A |
| 17 | 35.2 | 11.7 | 352 | 8 | US-11-266-748A-257564 | Sequence 257564, |
| 18 | 35.2 | 11.7 | 352 | 8 | US-11-266-748A-278222 | Sequence 278222, |
| 19 | 35.2 | 11.7 | 352 | 8 | US-11-266-748A-318081 | Sequence 318081, |
| 20 | 35.2 | 11.7 | 1421559 | 8 | US-11-266-748A-28208 | Sequence 28208, A |
| 21 | 34.6 | 11.5 | 633 | 6 | US-10-471-571A-3251 | Sequence 3251, Ap |
| 22 | 33.6 | 11.2 | 1108 | 8 | US-11-266-748A-353636 | Sequence 353636, |
| 23 | 33.6 | 11.2 | 1108 | 8 | US-11-266-748A-437015 | Sequence 437015, |
| 24 | 33.6 | 11.2 | 1908 | 8 | US-11-266-748A-188843 | Sequence 188843, |

| | | | | | |
|------|------|---------|---|-----------------------|-------------------|
| 33.6 | 11.2 | 10389 | 8 | US-11-266-748A-30896 | Sequence 30896, A |
| 33.4 | 11.1 | 1000 | 8 | US-11-266-748A-200842 | Sequence 200842, |
| 33.4 | 11.1 | 1000 | 8 | US-11-266-748A-201296 | Sequence 201296, |
| 33.4 | 11.1 | 1278 | 8 | US-11-217-529-81496 | Sequence 81496, A |
| 33.2 | 11.1 | 1146 | 8 | US-11-266-748A-77181 | Sequence 77181, A |
| 33.2 | 11.1 | 1146 | 8 | US-11-266-748A-129992 | Sequence 129992, |
| 32.8 | 10.9 | 594 | 8 | US-11-217-529-5685 | Sequence 5685, Ap |
| 32.8 | 10.9 | 109669 | 7 | US-11-021-837-41 | Sequence 41, Appl |
| 32.8 | 10.9 | 111597 | 8 | US-11-266-748A-22665 | Sequence 22665, A |
| 32.8 | 10.9 | 247877 | 8 | US-11-266-748A-22663 | Sequence 22663, A |
| 32.8 | 10.9 | 1071650 | 8 | US-11-266-748A-22664 | Sequence 22664, A |
| 32.6 | 10.9 | 4857 | 6 | US-10-517-441-424 | Sequence 424, App |
| 32.4 | 10.8 | 1000 | 8 | US-11-266-748A-224134 | Sequence 224134, |
| 32.4 | 10.8 | 1000 | 8 | US-11-266-748A-292677 | Sequence 292677, |
| 32.4 | 10.8 | 1000 | 8 | US-11-266-748A-344106 | Sequence 344106, |
| 32.4 | 10.8 | 1000 | 8 | US-11-266-748A-404470 | Sequence 404470, |
| 32.4 | 10.8 | 1000 | 8 | US-11-266-748A-45516 | Sequence 475516, |
| 32.4 | 10.8 | 1039 | 8 | US-11-266-748A-71950 | Sequence 71950, A |
| 32.4 | 10.8 | 1039 | 8 | US-11-266-748A-124761 | Sequence 124761, |
| 32.4 | 10.8 | 8391 | 8 | US-11-266-748A-24100 | Sequence 24100, A |
| 32.4 | 10.8 | 31857 | 8 | US-11-217-529-166179 | Sequence 166179, |
| 32.2 | 10.7 | 1218 | 8 | US-11-266-748A-22572 | Sequence 22572, A |
| 32.2 | 10.7 | 1218 | 8 | US-11-266-748A-82454 | Sequence 82454, A |
| 32.2 | 10.7 | 1218 | 8 | US-11-266-748A-111321 | Sequence 111321, |
| 32.2 | 10.7 | 4388 | 8 | US-11-266-748A-135265 | Sequence 135265, |
| 32.2 | 10.7 | 4388 | 8 | US-11-266-748A-30890 | Sequence 30890, A |
| 32.2 | 10.7 | 26683 | 7 | US-11-021-837-26 | Sequence 26, Appl |
| 32 | 10.7 | 1113 | 8 | US-11-217-529-82444 | Sequence 82444, A |
| 32 | 10.7 | 1924 | 8 | US-11-266-748A-73868 | Sequence 73868, A |
| 32 | 10.7 | 1924 | 8 | US-11-266-748A-126679 | Sequence 126679, |
| 32 | 10.7 | 7126 | 8 | US-11-266-748A-58818 | Sequence 58818, A |
| 32 | 10.7 | 7326 | 8 | US-11-266-748A-28338 | Sequence 28338, A |
| 32 | 10.7 | 140167 | 8 | US-11-266-748A-60770 | Sequence 60770, A |
| 31.8 | 10.6 | 527 | 8 | US-11-266-748A-307049 | Sequence 307049, |
| 31.8 | 10.6 | 637 | 8 | US-11-216-545-3893 | Sequence 3893, Ap |
| 31.8 | 10.6 | 693 | 8 | US-11-266-748A-35109 | Sequence 35109, A |
| 31.8 | 10.6 | 932 | 8 | US-11-266-748A-19166 | Sequence 19166, A |
| 31.8 | 10.6 | 4722 | 8 | US-11-266-748A-28340 | Sequence 28340, A |
| 31.6 | 10.5 | 458 | 8 | US-11-266-748A-92792 | Sequence 92792, A |
| 31.6 | 10.5 | 458 | 8 | US-11-266-748A-145603 | Sequence 145603, |
| 31.6 | 10.5 | 1000 | 8 | US-11-266-748A-281058 | Sequence 281058, |
| 31.6 | 10.5 | 1000 | 8 | US-11-266-748A-307698 | Sequence 307698, |
| 31.6 | 10.5 | 1000 | 8 | US-11-266-748A-398345 | Sequence 398345, |
| 31.6 | 10.5 | 1000 | 8 | US-11-266-748A-469391 | Sequence 469391, |
| 31.6 | 10.5 | 1219 | 6 | US-10-509-131-27 | Sequence 27, Appl |
| 31.6 | 10.5 | 1409 | 7 | US-11-218-305-23785 | Sequence 23785, A |
| 31.6 | 10.5 | 1807 | 8 | US-11-216-545-6502 | Sequence 6502, Ap |
| 31.6 | 10.5 | 5275 | 8 | US-11-266-748A-61192 | Sequence 61192, A |
| 31.4 | 10.5 | 654 | 8 | US-11-266-748A-412734 | Sequence 412734, |
| 31.4 | 10.5 | 663 | 8 | US-11-266-748A-49205 | Sequence 49205, A |
| 31.4 | 10.5 | 663 | 8 | US-11-266-748A-209620 | Sequence 209620, |
| 31.4 | 10.5 | 663 | 8 | US-11-266-748A-234110 | Sequence 234110, |
| 31.4 | 10.5 | 892 | 8 | US-11-266-748A-182511 | Sequence 182511, |
| 31.4 | 10.5 | 892 | 8 | US-11-266-748A-243056 | Sequence 243056, |
| 31.4 | 10.5 | 5189 | 8 | US-11-266-748A-27190 | Sequence 27190, A |
| 31.4 | 10.5 | 5189 | 8 | US-11-266-748A-30394 | Sequence 30394, A |
| 31.4 | 10.5 | 5967 | 8 | US-11-266-748A-32555 | Sequence 32555, A |
| 31.4 | 10.5 | 8622 | 8 | US-11-319-952-21 | Sequence 21, Appl |
| 31.2 | 10.4 | 901 | 8 | US-11-266-748A-354433 | Sequence 354433, |
| 31.2 | 10.4 | 901 | 8 | US-11-266-748A-437812 | Sequence 437812, |
| 31.2 | 10.4 | 913 | 8 | US-11-266-748A-8033 | Sequence 8033, Ap |
| 31.2 | 10.4 | 1194 | 8 | US-11-217-529-81713 | Sequence 81713, A |
| 31.2 | 10.4 | 1563 | 8 | US-11-217-529-75623 | Sequence 75623, A |
| 31.2 | 10.4 | 2503 | 8 | US-11-266-748A-61671 | Sequence 61671, A |
| 31.2 | 10.4 | 55266 | 8 | US-11-266-748A-24926 | Sequence 24926, A |
| 31 | 10.3 | 846 | 8 | US-11-266-748A-296367 | Sequence 296367, |
| 31 | 10.3 | 1000 | 8 | US-11-266-748A-198313 | Sequence 198313, |
| 31 | 10.3 | 1427 | 8 | US-11-266-748A-193857 | Sequence 193857, |
| 31 | 10.3 | 1585 | 8 | US-11-266-748A-181087 | Sequence 181087, |
| 31 | 10.3 | 5405 | 8 | US-11-266-748A-29238 | Sequence 29238, A |
| 31 | 10.3 | 150000 | 8 | US-11-266-748A-23591 | Sequence 23591, A |
| 31 | 10.3 | 684973 | 8 | US-11-266-748A-32013 | Sequence 32013, A |
| 30.8 | 10.3 | 9001 | 6 | US-10-517-441-790 | Sequence 790, App |

```
c 98 30.8 10.3 1421559 8 US-11-266-748A-28208
c 99 30.6 10.2 296 8 US-11-266-748A-419450,
c 100 30.6 10.2 984 8 US-11-266-748A-259017,
c 101 30.6 10.2 984 8 US-11-266-748A-319534,
c 102 30.6 10.2 1000 8 US-11-266-748A-224369,
c 103 30.6 10.2 1000 8 US-11-266-748A-281385,
c 104 30.6 10.2 1000 8 US-11-266-748A-293251,
c 105 30.6 10.2 1000 8 US-11-266-748A-308025,
c 106 30.6 10.2 1000 8 US-11-266-748A-344680,
c 107 30.6 10.2 1000 8 US-11-266-748A-405201,
c 108 30.6 10.2 1000 8 US-11-266-748A-476247,
c 109 30.6 10.2 1022 8 US-11-266-748A-371060,
c 110 30.6 10.2 1022 8 US-11-266-748A-454439,
c 111 30.6 10.2 1483 7 US-11-218-305-13433,
c 112 30.6 10.2 1590 6 US-10-471-571A-2589,
c 113 30.6 10.2 2083 8 US-11-266-748A-185657,
c 114 30.6 10.2 2083 8 US-11-266-748A-241015,
c 115 30.6 10.2 2298 8 US-11-266-748A-26651,
c 116 30.6 10.2 2618 8 US-11-266-748A-26534,
c 117 30.6 10.2 2879 8 US-11-266-748A-22882,
c 118 30.6 10.2 183483 8 US-11-266-748A-58663,
c 119 30.4 10.1 727 8 US-11-266-748A-212002,
c 120 30.4 10.1 727 8 US-11-266-748A-235525,
c 121 30.4 10.1 779 8 US-11-266-748A-54657,
c 122 30.4 10.1 779 8 US-11-266-748A-218259,
c 123 30.4 10.1 959 8 US-11-266-748A-251091,
c 124 30.4 10.1 1000 8 US-11-266-748A-311608,
c 125 30.4 10.1 1000 8 US-11-266-748A-116353,
c 126 30.4 10.1 1000 8 US-11-266-748A-158517,
c 127 30.4 10.1 1000 8 US-11-266-748A-397281,
c 128 30.4 10.1 1000 8 US-11-266-748A-468327,
c 129 30.4 10.1 1182 8 US-11-266-748A-178335,
c 130 30.4 10.1 1553 7 US-11-266-748A-182159,
c 131 30.4 10.1 1555 7 US-11-218-305-14463,
c 132 30.4 10.1 3189 8 US-11-293-697-748,
c 133 30.4 10.1 8900 6 US-10-517-441-427,
c 134 30.4 10.1 8900 6 US-10-517-441-701,
c 135 30.4 10.1 38945 8 US-11-266-748A-24969,
c 136 30.4 10.1 150024 8 US-11-266-748A-60141,
c 137 30.2 10.1 479 8 US-11-266-748A-178820,
c 138 30.2 10.1 596 8 US-11-266-748A-51992,
c 139 30.2 10.1 620 8 US-11-266-748A-91245,
c 140 30.2 10.1 620 8 US-11-266-748A-144056,
c 141 30.2 10.1 920 8 US-11-266-748A-69209,
c 142 30.2 10.1 920 8 US-11-266-748A-105300,
c 143 30.2 10.1 920 8 US-11-266-748A-122020,
c 144 30.2 10.1 967 8 US-11-266-748A-353019,
c 145 30.2 10.1 967 8 US-11-266-748A-436398,
c 146 30.2 10.1 1000 8 US-11-266-748A-223747,
c 147 30.2 10.1 1000 8 US-11-266-748A-291799,
c 148 30.2 10.1 1000 8 US-11-266-748A-343228,
c 149 30.2 10.1 1000 8 US-11-266-748A-403335,
c 150 30.2 10.1 1000 8 US-11-266-748A-474381,

Sequence 28208, A
Sequence 419450,
Sequence 259017,
Sequence 319534,
Sequence 224369,
Sequence 281385,
Sequence 293251,
Sequence 308025,
Sequence 344680,
Sequence 405201,
Sequence 476247,
Sequence 371060,
Sequence 454439,
Sequence 14333, A
Sequence 2589, Ap
Sequence 185657,
Sequence 241015,
Sequence 26651, A
Sequence 26534, A
Sequence 22882, A
Sequence 58663, A
Sequence 212002,
Sequence 235525,
Sequence 54657, A
Sequence 218259,
Sequence 251091,
Sequence 311608,
Sequence 116353,
Sequence 158517,
Sequence 397281,
Sequence 468327,
Sequence 178335,
Sequence 182159,
Sequence 14463, A
Sequence 748, App
Sequence 427, App
Sequence 701, App
Sequence 24969, A
Sequence 60141, A
Sequence 178820,
Sequence 51992, A
Sequence 91245, A
Sequence 144056,
Sequence 69209, A
Sequence 105300,
Sequence 122020,
Sequence 353019,
Sequence 436398,
Sequence 223747,
Sequence 291799,
Sequence 343228,
Sequence 403335,
Sequence 474381,

; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 249055
; LENGTH: 863
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-11-266-748A-249055

Query Match 12.9%; Score 38.6; DB 8; Length 863;
Best Local Similarity 48.4%; Pred. No. 0.38;
Matches 107; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Qy 41 TTGAATATTTCTTCATTTGAAATAAACAGCTAACTCCCAAAATTTGCTACCTAACTAGGGGAG 100
Db 26 TTGACTTTTTTATTCAGGGAATAAAGTTTTTACTTTGGAAACAGATATTTTAGCATAGA 85
Qy 101 TTTATCATGAGGAAGTGTATTAATGCTTATTTTCAATCACTATATAGCTGTAGTAGTGCATTTCT 160
Db 86 CTTAGCATAGAGCAAACTAAATATGATTAGTATTTCCAAAGGTACAAATCATATTTGGCC 145
Qy 161 AACATGCTGTATAATAACCGCAAACTTCAATCACTATATAGCTGTAGTAGTGCATTTCT 220
Db 146 AAATCCTTCAAAAGCAAGTACCCTACTATACAGACAGATCTGGGTGACGATCGCTGCT 205
Qy 221 GCAAGGATCCAGAGTAACAGTATTTTGGAAATGCAATGT 261
Db 206 TATTGCAACAGTGAATCTCTGCGACCTTAGAATGTCCTGCT 246

RESULT 2
US-11-266-748A-201627/c
; Sequence 201627, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
```



```
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-481257

Query Match      12.6%; Score 37.8; DB 8; Length 1000;
Best Local Similarity 57.0%; Pred. No. 0.65; Indels 0; Gaps 0;
Matches 69; Conservative 0; Mismatches 52;

QY      8 ACAGGCTAAATGCTAGTACGATGAGACTTGAAGACTTGAATATTTCTTCATTGAAATAAAC 67
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      270 ACTGCCCTACAACAAGGCGAGTTCTCTGAACAACACTGATTTATTTTCATTATCAATCA 329

QY      68 AGCTAACTCCCAATTTGTACCTAACTAGGGAGTTTATCATGAGAAATGTTTAAATGCT 127
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      330 ATTATATTTTCAAAATTTATACCTTAAAAATGTTAGTTGGTAAATGAACATAGTAAGT 389

QY      128 T 128
      |
Db      390 T 390

RESULT 7
US-11-266-748A-23772/c
/ Sequence 23772, Application US/11266748A
/ Publication No. US20060134663A1
/ GENERAL INFORMATION:
/ APPLICANT: Johnston, Paul
/ APPLICANT: Mulligan, Karl
/ TITLE OF INVENTION: Transcriptome Microarray Technology and
/ TITLE OF INVENTION: Methods of Using the Same
/ FILE REFERENCE: 55815-0102 (319189)
/ CURRENT APPLICATION NUMBER: US/11/266,748A
/ CURRENT FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/662,276
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 483996
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 23772
/ LENGTH: 6857
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-23772

Query Match      12.3%; Score 37; DB 8; Length 6857;
Best Local Similarity 55.9%; Pred. No. 1.7; Indels 1; Gaps 1;
Matches 90; Conservative 0; Mismatches 70;

QY      41 TTGAATATTTCTTCATTGTTGAAATTAACAGCTAACTCCCAATTTGTACCTAAGGGAG 100
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2234 TTGAATATTTTAAAGCTTAAATTTCTGCTGGTTTGCAAAATTTCTTTTAACTATATC 2175

QY      101 TTATCATGAA-GAAATGTTTAAATGCTTAATTTTCAATAAGAAAGTAAACAGAGTGTG 159
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2174 TATATTATTAACATGTAAGTATTTTTCCTGCTGATTATGAAATATAAAGTAACACAGGGTTA 2115

QY      160 CAACATGCTGTAAATAACCCGACAAACTTCAATCACTATA 200
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2114 CAGTTGTGCAAAACAAACAAACATACCTAAATTAACAATA 2074
```

```
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-390539

Query Match      12.6%; Score 37.8; DB 8; Length 1000;
Best Local Similarity 57.0%; Pred. No. 0.65; Indels 0; Gaps 0;
Matches 69; Conservative 0; Mismatches 52;

QY      8 ACAGGCTAAATGCTAGTACGATGAGACTTGAAGACTTGAATATTTCTTCATTGAAATAAAC 67
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      731 ACTGCCCTACAACAAGGCGAGTTCTCTGAACAACACTGATTTATTTTCATTATCAATCA 672

QY      68 AGCTAACTCCCAATTTGTACCTAACTAGGGAGTTTATCATGAGAAATGTTTAAATGCT 127
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      671 ATTATATTTTCAAAATTTATACCTTAAAAATGTTAGTTGGTAAATGAACATAGTAAGT 612

QY      128 T 128
      |
Db      611 T 611

RESULT 6
US-11-266-748A-481257
/ Sequence 481257, Application US/11266748A
/ Publication No. US20060134663A1
/ GENERAL INFORMATION:
/ APPLICANT: Johnston, Paul
/ APPLICANT: Mulligan, Karl
/ TITLE OF INVENTION: Transcriptome Microarray Technology and
/ TITLE OF INVENTION: Methods of Using the Same
/ FILE REFERENCE: 55815-0102 (319189)
/ CURRENT APPLICATION NUMBER: US/11/266,748A
/ CURRENT FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/662,276
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 483996
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 481257
/ LENGTH: 1000
```

```

; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 360446
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (12)..(12)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (66)..(66)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (73)..(73)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131)..(131)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (376)..(376)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (394)..(394)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-360446

Query Match      12.1%; Score 36.4; DB 8; Length 415;
Best Local Similarity 51.2%; Pred. No. 1.3;
Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 46 ATATCTTCATTGGAATAAAGACCTAACTCCCAAAATGTACCTAGGGAGTTTAT 105
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 ATATAAATTCAGTCACAAATAATGTAATGGAAGTAACTCCAGTTAGATTGAATTAAT 241
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 106 CATCAAGAATAATGTTTAAATGCTTAATTTTCAATAAAGAAAGTAACCAAGTGTTCACAT 165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 AAAATCCAGCTGTTTTTAACTATATATATATATATATATATATATATATATATATATATAT 301
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 166 GCTGTTAAATAACCCGACAACTTCAATCACTATAGCTGTAGTAGA 211
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 302 TTTATAAAGGAGAGTATAAAGATATGCCAGTATACATTAAGTACTAGA 347
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
US-11-266-748A-443825/c
; Sequence 443825, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl

```

```

; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 16908
; LENGTH: 692
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-16908

Query Match      12.3%; Score 36.8; DB 8; Length 692;
Best Local Similarity 50.4%; Pred. No. 1.1;
Matches 121; Conservative 0; Mismatches 112; Indels 7; Gaps 1;

Qy 44 AAATATCTTCATTGGAATAAAGACCTAACTCCCAAAATGTACCTAGGGAGTTT 103
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 248 AGATATTTTAAATTAACATACACATTTTCTTTTAAAGGTCTATCTATTAGGCCTT 307
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 104 ATCATGAAGAATGTTTAAATGCTTAATTTTCAATAAAGAAAGTAACCAAGTGTTCGAC 163
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 TCCCTGCTGGTAACCTTAAATGATTTATTTGTGAATTTTAACTAA-----GTTTCATC 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 164 ATGCTGTTTAAATAACCCGACAACTTCAATCACTATAGCTGTAGTAGAGTGCATTCTGCA 223
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 AAGATCTTAAAGAGAGAGAGAAATAAAGATATTATTATATGTCAGTGCCCACTTTG 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 224 AGGATCCAGAGTAACCAAGTATTTTGGAAATGCAATGTTGAACCCAGCATCTAATATATC 283
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 AAGGTAAACAGTAGTAGTAATCTTTTGGAGAAATTTCTAAAACTACTTTTAAAGAGGC 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
US-11-266-748A-360446
; Sequence 360446, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03

```

; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 443825
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (22)..(22)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40)..(40)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (285)..(285)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (343)..(343)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (350)..(350)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (404)..(404)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-443825

Query Match 12.1%; Score 36.4; DB 8; Length 415;
Best Local Similarity 51.2%; Pred. No. 1.3;
Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
Qy 46 ATATCTCTTTTGAATAAAGCTTAACCTCCCAAAATTTGACCTAAGGGAGTTTAT 105
Db 234 ATATAATCAGTCACAACAATGTAATGGCTAAAGTTACCCAGTTAGATTGAATTAAT 175
Qy 106 CATGAAGAAATGTTTAAATGCTTAATTTCAAAATGAAGTAACACAGAGTGTTCACACAT 165
Db 174 AAAATCCAGCTGTTTTTAATCTATATTTTAAATATGAAGGTACAAAAGGTTGAAAGTT 115
Qy 166 GCTGTTAAATACCCGACAACCTCAATCACTATAGCTAGTAGAGA 211
Db 114 TTTTAAAGGAGAGTGAATAAGATATGCCAGTATACATACTAGA 69

RESULT 11
US-11-266-748A-362773
; Sequence 362773, Application US/11266748A

; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 362773
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-362773

Query Match 11.9%; Score 35.6; DB 8; Length 578;
Best Local Similarity 50.6%; Pred. No. 2.2;
Matches 86; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
Qy 36 AAGACTTGAATATTTCTTCAATTTGAATAAAGCTTAACCTCCCAAAATTTGACCTAAGT 95
Db 409 AGGACATGGAGTCTACTGTATTTAAATATAGAGCTCATCTCAAAATTTCTATATGTAAG 468
Qy 96 GCGAGTTTATCATGAAGAATGTTTAAATGCTTTAAATTTCAATAAAGAGTAACCAAGT 155
Db 469 AGCAATTTTCTTTGAAAACATGATATATGGATAATAATAATTAATTAATTAATGTTT 528
Qy 156 GTTGCAACATGCTGTTAAATAACCCGACAACCTTCAATCACTACTATAGCTGT 205
Db 529 GAAATATCAGACATAGNAATCATAGCATCTATTAAACAACTTTGCTTT 578

RESULT 12
US-11-266-748A-446152/c
; Sequence 446152, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9


```

; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 5815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31158
; LENGTH: 2205
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-31158

Query Match      11.9%; Score 35.6; DB 8; Length 2205;
Best Local Similarity 50.6%; Pred. No. 3.1;
Matches 86; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 36 AAGCTTGAATATCTTCTTCAATTAACAGCTAATCCCAAAATGTTACCTAAGT 95
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2008 AGGACATGGAGTCTACTGTATTTAAATATAGAGCTCATCTCAAAATCTATGTAAAG 2067
QY 96 GGGAGTTTATCATGAAGAATGTTTAAATGCTTAAATTTTCAAATAAGAGTAAACAGAGT 155
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2068 AGCATTTCTTTGAAAACATGATATATGGATAAATAATTAATCTACTTCTTAAATGTTT 2127
QY 156 GTTCCAAATGCTGTTTAAATTAACCCGACAACTTCAATCACTATAGCTGT 205
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2128 GAAATATCAGACATAGAAATCATAGCATCTATTAAACAACTTTGCCTTT 2177

RESULT 16
US-11-266-748A-58551/c
; Sequence 58551, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 5815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31158
; LENGTH: 2205
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (50)..(50)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:

```

```

; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 58551
; LENGTH: 130263
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-58551

Query Match      11.9%; Score 35.6; DB 8; Length 130263;
Best Local Similarity 49.5%; Pred. No. 8;
Matches 92; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 24 GTGACTACGATGAAGACTTGAATAATTTCTTCAATTAACAGCTAATCCCAAAAT 83
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
36114 GTGAATAACAGGTACACTGTCACTCATATAAATCCTGAATACCGACCTAATCTGTCAGTA 36055
QY 84 GTACCTAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTCAAAATAGA 143
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
36054 TGAGTGGGAAATGGAAGGATAGTTATGTGAGACAGATTAACTTCAACTTCCAAATAGA 35995
QY 144 AGTAACACAGAGTGTTCGAACATCTGTTAAATTAACCCGACAACTTCAATCACTATAGCT 203
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
35994 AGTCCACAGAGATATCCAAAGTGAAGAGCTAAGAAATAGCAGTATAAAACATTAAAAAT 35935
QY 204 GTAGTA 209
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
35934 GAAATA 35929

RESULT 17
US-11-266-748A-257564/c
; Sequence 257564, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 5815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 257564
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (50)..(50)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:

```

```

; NAME/KEY: misc_feature
; LOCATION: (67)..(67)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (74)..(74)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117)..(117)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (161)..(161)
; OTHER INFORMATION: n is a, c, g, or t
; US-11-266-748A-257564

Query Match      11.7%; Score 35.2; DB 8; Length 352;
Best Local Similarity 49.4%; Pred. No. 2.5;
Matches 88; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 59 GAAATAAACAGCTAACTCCCAAAATGTACCTAACTAGGGAGTTTATCATGAAGAAATGT 118
Db 272 GAACAAATATCAGTAATCCTTTGTTCTTAACAAAAATTCATAATTTATTATACATTT 213
Qy 119 TTAATGCTTAATTTTCAAATAAGAGTAACAGAGTGTGCAACATGCTGTTAAATAAC 178
Db 212 TAAATATTATTGTTTCAAATGTTTAAAGATGTTAGTGGGCAATAAATCATAAAGNGATACAAC 153
Qy 179 CCGCAAACTTCAATCACTATAGCTGTAGTAGAGTGCAATTCGCAAGATCCAGAGT 236
Db 152 CAGTCGAAATTAGACTTTTAAAGATTTACTCTCAANGACATTTGCAACAACTCCAAAGT 95

RESULT 18
US-11-266-748A-278222/c
; Sequence 278222, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 278222
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (50)..(50)
; OTHER INFORMATION: n is a, c, g, or t

; NAME/KEY: misc_feature
; LOCATION: (67)..(67)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (74)..(74)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117)..(117)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (161)..(161)
; OTHER INFORMATION: n is a, c, g, or t
; US-11-266-748A-278222

Query Match      11.7%; Score 35.2; DB 8; Length 352;
Best Local Similarity 49.4%; Pred. No. 2.5;
Matches 88; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 59 GAAATAAACAGCTAACTCCCAAAATGTACCTAACTAGGGAGTTTATCATGAAGAAATGT 118
Db 272 GAACAAATATCAGTAATCCTTTGTTCTTAACAAAAATTCATAATTTATTATACATTT 213
Qy 119 TTAATGCTTAATTTTCAAATAAGAGTAACAGAGTGTGCAACATGCTGTTAAATAAC 178
Db 212 TAAATATTATTGTTTCAAATGTTTAAAGATGTTAGTGGGCAATAAATCATAAAGNGATACAAC 153
Qy 179 CCGCAAACTTCAATCACTATAGCTGTAGTAGAGTGCAATTCGCAAGATCCAGAGT 236
Db 152 CAGTCGAAATTAGACTTTTAAAGATTTACTCTCAANGACATTTGCAACAACTCCAAAGT 95

RESULT 19
US-11-266-748A-318081
; Sequence 318081, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 318081
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (192)..(192)

```

```
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (236)..(236)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (279)..(279)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (286)..(286)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (303)..(303)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-318081
```

```
Query Match 11.7%; Score 35.2; DB 8; Length 352;
Best Local Similarity 49.4%; Pred. No. 2.5;
Matches 88; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 59 GAAATAACAGCTAACTCCCAAAATGTACCTTAAGTGGGAGTTTATCATGAAGAAATGT 118
Db 81 GAAACAATATCAGTAATCTCTTTGTTCTTAACAAAAATTCATAATTTATATACATT 140

Qy 119 TTAATGCTTTAAATTTTCAAAATAAGAGTAACAGAGTGTTCACATCGTCTTTAAATPAAC 178
Db 141 TAAATATTATTATGTTTCAAAATGTTGTAGTGGGCAATAAATCATAAAGNGATACAAC 200

Qy 179 CCGACAAACTTCAATCACTATAGCTGTAGTAGATGCAATTCGCAAGGATCCAGAGT 236
Db 201 CAGTCGAATTAGAGCTTTTAAAGATTTACTCTCAANGACATTCGCAACAACTCCAAAGT 258
```

RESULT 20

```
US-11-266-748A-28208
; Sequence 28208, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28208
; LENGTH: 1421559
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (414394)..(414394)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (836909)..(836909)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1022504)..(1022504)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1041105)..(1041105)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1122667)..(1122667)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-28208

Query Match 11.7%; Score 35.2; DB 8; Length 1421559;
Best Local Similarity 53.7%; Pred. No. 18;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 15 AAATGCTAAGTGAAGTACGATGAAGACTTGAAATATTCTTCATTTGAAATAAACAGCTAAC 74
Db 289724 AAAAAAGAAAAACAAAAGAGATCAATTGAATCAATTTGAATGATATATAAGAAATA 289783

Qy 75 TCCCAAAATGTACTTAACCTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTT 134
Db 289784 GCCAAATAGTAGGTGATAGGAGAAATATCAACAAGGTGCATGGAGTAATTTTCCAAATTT 289843

Qy 135 CAAATAAGAGTAACC 150
Db 289844 TGAGGCTTAGTAACC 289859
```

RESULT 21

```
US-10-471-571A-3251/c
; Sequence 3251, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3251
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-3251

Query Match 11.5%; Score 34.6; DB 6; Length 633;
Best Local Similarity 56.6%; Pred. No. 4.2;
Matches 64; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 105 TCATGAAGAAATGTTTAAATGCTTAATTTTCAATAAGAAAGTACACGAGTGTTCACACA 164
Db 389 TTATTATTATAAATTTCCAACTTAGCTTAGATTTCAGAAATAAAAAAGTATTAAACA 330

Qy 165 TGCTGTTAAATAACCCGACAAACTTCAATCACTATAGCTGTAGTAGAGTGCAT 217
Db 329 CCCAGAAATATTGCACAACATATTACATCACCATGTTTCTCTAATTGAAT 277
```

RESULT 22

```
US-11-266-748A-353636/c
; Sequence 353636, Application US/11266748A
```

Publication No. US20060134663A1

GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
METHODS OF USING THE SAME
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 353636
LENGTH: 1108
TYPE: DNA

ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1060)..
OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-353636

Query Match 11.2%; Score 33.6; DB 8; Length 1108;
Best Local Similarity 46.6%; Pred. No. 9;
Matches 108; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 29 TACGATGAAGACTTGAATAATTTCTTCAATTTGAAATAAACAGCTAACTCCCAAAATGTACC 88
DB 257 TGCATAGTAGTTTCGATTAACTGATAATTTGGGTTCCCACTGTAAGATATAAAC 198
QY 89 TAACTAGGGAGTTTATCATGAGAAATGTTTAAATGCTTAAATTTCAAATAAGAGTAA 148
DB 197 AGAATGGAGAAATTAATGGAGAGTAACCTTTTCATAGCTGTATTATATAAAGGGTGCACAC 138
QY 149 CCAGAGTGTTCACATGCTGTAAATAAACCCGACAACTTCAATCACTATAGCTGTAGT 208
DB 137 ATTTGACAGCTCAGACACTTTGATCAAGAGCACTACTAGCAAGTGTCAAGTGTGGG 78
QY 209 AGAGTGCAATCTGCAAGGATCCAGAGTAACCCAGATTAATTTGGAAATGCAATG 260
DB 77 CAACGTCTTCTTCAGGCTCCAGAAAGCACTTATTCTTGGTGAAGAAAG 26

RESULT 23

US-11-266-748A-437015
Sequence 437015, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
METHODS OF USING THE SAME
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2

PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 437015
LENGTH: 1108
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (49)..
OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-437015

Query Match 11.2%; Score 33.6; DB 8; Length 1108;
Best Local Similarity 46.6%; Pred. No. 9;
Matches 108; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 29 TACGATGAAGACTTGAATAATTTCTTCAATTTGAAATAAACAGCTAACTCCCAAAATGTACC 88
DB 852 TGCATAGTAGTTTCGATTAACTGATAATTTGGGTTCCCACTGTAAGATATAAAC 911
QY 89 TAACTAGGGAGTTTATCATGAGAAATGTTTAAATGCTTAAATTTCAAATAAGAGTAA 148
DB 912 AGAATGGAGAAATTAATGGAGAGTAACCTTTTCATAGCTGTATTATATAAAGGGTGCACAC 971
QY 149 CCAGAGTGTTCACATGCTGTAAATAAACCCGACAACTTCAATCACTATAGCTGTAGT 208
DB 972 ATTTGACAGCTCAGACACTTTGATCAAGAGCACTACTAGCAAGTGTCAAGTGTGGG 1031
QY 209 AGAGTGCAATCTGCAAGGATCCAGAGTAACCCAGATTAATTTGGAAATGCAATG 260
DB 1032 CAACGTCTTCTTCAGGCTCCAGAAAGCACTTATTCTTGGTGAAGAAAG 1083

RESULT 24

US-11-266-748A-188843/c
Sequence 188843, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
METHODS OF USING THE SAME
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2

```
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 188843
; LENGTH: 1908
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1676)..(1684)
; OTHER INFORMATION: n is a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (1790)..(1819)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-188843

Query Match      11.2%; Score 33.6; DB 8; Length 1908;
Best Local Similarity 46.6%; Pred. No. 10;
Matches 108; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

29  TACGATGAAGACTTGAAATATCTTCATTTGGAATAAACAGCTAACTCCCAAAATGTACC 88
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
877  TCGAATGTAGTTTCGATTAACTGATTAATTTGGAATTTGGGTCCTCAACTGTAGATATAAC 818
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
89  TAACTAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAATTTTCAAATAAGAAGTAA 148
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
817  AGAATGGAGAAATTAATGGAGAAGTAATCTTTTCATAGCTGTATTATAAAGGTTGCACAC 758
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149  CCAGAGTGTCCACATGCTGTTAAATAACCCGACAACTTCAATCACTATAGCTGTAGT 208
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
757  ATTTGACAGCCTCAGACACTCTTGATCAAGAGCCTACTAGCAAGTGTCAAAGTGTGGG 698
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
209  AGAGTGCATCTCGAAGGATCCAGAGTAACAGGATTTTGGAAATGCAATG 260
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
697  CAACTGCTCTTCGAGGCTCCAGAAAGAACCTTTATCTTGGTCAAGGAAAG 646

RESULT 25
US-11-266-748A-30896/c
; Sequence 30896, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Patrick
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 30896
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-200842

Query Match      11.1%; Score 33.4; DB 8; Length 1000;
Best Local Similarity 51.0%; Pred. No. 10;
Matches 79; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

60  AATAACAGCTAACTCCCAAAATGTACTTAACCTAGGGAGTTTATCATGAAGAATGTT 119
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
763  AGAATAATATCTAGACAGATAAATATGCTGTACTGGGAGTTTATTAATGAAGAATGTA 704
Qy  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120  TAAATGCTTAATTTTCAAATAAGAAGTAACAGAGTGTTCACACATGCTGTAAATAACC 179
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


| | | | |
|----|------|------------------------------------------------------------------|------|
| Db | 1127 | TTCAATGAACATACATAGCACTAAATATGAGAGCTTCCAAAATGTACACGCAATTGGAT | 1068 |
| Qy | 101 | TTTATCATGAAGAAATGCTTTAAATGCTTAAATTTTCAAAATAGAAGTAACACAGAGTGTTC | 160 |
| Db | 1067 | CTAACTAACTAGGAGAGCTTCAGAAATTTTAAATTTAAAGATAAAAAGTGTTCACCTTGTGAAA | 1008 |
| Qy | 161 | AACATGCTGTTAAATAACCCGACAAACTCTCAATCACTATAGCTGTA | 206 |
| Db | 1007 | AATTACTATTAAATATATTTTAAATTAAGAAATTTTAAACATATA | 962 |

RESULT 30

US-11-266-748A-12992
; Sequence 12992, Application US/11266748A
; Publication No. US20060134663A1

GENERAL INFORMATION:

APPLICANT: Harkin, Paul
 APPLICANT: Johnston, Patrick
 APPLICANT: Mulligan, Karl
 TITLE OF INVENTION: Transcriptome Microarray Technology and
 TITLE OF INVENTION: Methods of Using the Same
 FILE REFERENCE: 5815-0102 (319189)
 CURRENT APPLICATION NUMBER: US/11/266,748A
 CURRENT FILING DATE: 2005-11-03
 PRIOR APPLICATION NUMBER: EP 04105479.2
 PRIOR FILING DATE: 2004-11-03
 PRIOR APPLICATION NUMBER: EP 04105482.6
 PRIOR FILING DATE: 2004-11-03
 PRIOR APPLICATION NUMBER: EP 04105483.4
 PRIOR FILING DATE: 2004-11-03
 PRIOR APPLICATION NUMBER: EP 04105507.0
 PRIOR FILING DATE: 2004-11-03
 PRIOR APPLICATION NUMBER: EP 04105485.9
 PRIOR FILING DATE: 2004-11-03
 PRIOR APPLICATION NUMBER: EP 04105484.2
 PRIOR FILING DATE: 2004-11-03
 PRIOR APPLICATION NUMBER: US 60/662,276
 PRIOR FILING DATE: 2005-03-14
 PRIOR APPLICATION NUMBER: US 60/700,293
 PRIOR FILING DATE: 2005-07-18
 NUMBER OF SEQ ID NOS: 48396
 SOFTWARE: PatentIn version 3.3
 SEQ ID NO 129992

| | Query Match | 11.1%; | Score 33.2; | DB 8; | Length 1146; |
|----|-----------------------|----------------------------------------------------------------|----------------|-----------|--------------|
| | Best Local Similarity | 50.0%; | Pred. No. 12; | | |
| | Matches 83; | Conservative 0; | Mismatches 83; | Indels 0; | Gaps 0; |
| Qy | 41 | TTGAAATATCTTTCATTGAAATAAACAGCTAACTCCCAAAATTCFACCTAACTAGGGGAG | 100 | | |
| Db | 20 | TTCAATGAAACTACATAGCACTAAATGAGAGCTTCCCAAAATGTAAACGACAAATTGGAT | 79 | | |
| Qy | 101 | TTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAAATAAGAAGTAACCAAGAGTGTTGC | 160 | | |
| Db | 80 | CTAACTAACTAGGGAGCTTCAGAATTTTAATTTAAAGNATAAAGTGTTCACTTGTGAA | 139 | | |
| Qy | 161 | AACATGCTGTTTAAATAACCCGACAAACTTCAATCACTATAGCTGTA | 206 | | |
| Db | 140 | AATTTACTATTAAATAATTTTTTAAATTAAGAATTTTTTAAACTATA | 185 | | |

RESULT 31

US-11-217-529-5685
; Sequence 5685, Application US/11217529
; Publication No. US20060099612A1

GENERAL INFORMATION:

APPLICANT: SUNTORY LIMITED
APPLICANT: NAKAO, YOSHIHIRO
APPLICANT: NAKAMURA, NORIHISA

```

; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 8-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; PRIOR FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5685
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
;
US-11-217-529-5685

Query Match          10.9%;   Score 32.8;   DB 8;   Length 594;
Best Local Similarity 56.5%;   Pred. No. 13;
Matches 61;   Conservative 0;   Mismatches 47;   Indels 0;   Gaps 0;

Qy      81  ATTGTACCTAACTAGGGGAGTGTTCATCATGAAGAAATGTTTAAATGCTTAAATTTCAAATA 140
      |||||
Db      206  ACTGGGACTATCAACAAGATTTCAACCAAGAAATGTTTCAGCACATAAATTCACGAAGA 265
      |||||

Qy      141  AGAAGTAACCAGAGTGTGTGCAACATGCTGTTTAAATAACCCGACAAACT 188
      |||||

Db      266  TAAATCATCTGAGAGTAAACGGCATAAAGAAATAATCTCGAAAAACT 313
      |||||

```

RESULT 32

US-11-021-837-41
; Sequence 41, Application US/11021837
; Publication No. US20060140972A1

: GENERAL INFORMATION:

```

/ APPLICANT: Alm, Richard
/ APPLICANT: Manning, Paul
/ APPLICANT: McLaughlin, Robert
/ APPLICANT: McCormack, Kathleen
/ TITLE OF INVENTION: Staphylococcus Saprophyticus
/ FILE REFERENCE: 101324-US
/ CURRENT APPLICATION NUMBER: US/11/021,837
/ CURRENT FILING DATE: 2004-12-23
/ PRIOR APPLICATION NUMBER: US 60/533534
/ PRIOR FILING DATE: 2003-12-31
/ PRIOR APPLICATION NUMBER: US 60/600680
/ PRIOR FILING DATE: 2003-08-31
/ NUMBER OF SEQ ID NOS: 63
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 1
/ LENGTH: 109669
/ TYPE: DNA
/ ORGANISM: Staphylococcus saprophyticus
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (46850)..(46850)
/ OTHER INFORMATION: n is a, c, g, or t
US-11-021-837-41

```

| Query Match | 10.9% | Score 32.8; | DB 7; | Length 109669; |
|-----------------------|----------------------------------------------------------------|----------------|-----------|----------------|
| Best Local Similarity | 52.1% | Pred. No. 44; | | |
| Matches 73; | Conservative 0; | Mismatches 67; | Indels 0; | Gaps 0; |
| Qy 2 | GCATTGACAGGCTAAATGCTAAGTGACTACGATGAAGACTTGAATAATCTTCATTGGAA | 61 | | |
| | | | | |
| Db 30576 | GCAAAATCAAATAAATTAACAACATCATTCAAATAATAGTTAGCTGTTTTTTTAAATTTGCG | 30635 | | |
| | | | | |
| Qy 62 | ATAACAGCTAACTCCCAAAATTGACCTTAACCTAGGGGAGTTTTATCATGAAGAAATGTTTA | 121 | | |
| | | | | |
| Db 30636 | AAAATGGCGAAATGGTTGGTTGGAAATAGTAAGCAGAGAGAAACAATAAAAAAATATTCA | 30695 | | |
| | | | | |
| Qy 122 | AATGCTTAATTTTCAAATAA | 141 | | |
| | | | | |

DB 30696 RAGGCTCTACGTTCTTAATAA 30715

RESULT 33

US-11-266-748A-22665/c

Sequence 22665, Application US/11266748A

Publication No. US20060134663A1

GENERAL INFORMATION:

APPLICANT: Harkin, Paul

APPLICANT: Johnston, Patrick

APPLICANT: Mulligan, Karl

TITLE OF INVENTION: Transcriptome Microarray Technology and Methods of Using the Same

FILE REFERENCE: 55815-0102 (319189)

CURRENT APPLICATION NUMBER: US/11/266,748A

CURRENT FILING DATE: 2005-11-03

PRIOR APPLICATION NUMBER: EP 04105479.2

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105482.6

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105483.4

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105485.9

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105507.0

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105485.9

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105484.2

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: US 60/662,276

PRIOR FILING DATE: 2005-03-14

PRIOR APPLICATION NUMBER: US 60/700,293

PRIOR FILING DATE: 2005-07-18

NUMBER OF SEQ ID NOS: 483996

SOFTWARE: PatentIn version 3.3

SEQ ID NO 22665

LENGTH: 111597

TYPE: DNA

ORGANISM: Homo Sapiens

US-11-266-748A-22665

Query Match 10.9%; Score 32.8; DB 8; Length 111597;

Best Local Similarity 51.4%; Pred. No. 44;

Matches 76; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 4 ATTGACAGGCTAAATGCTTAAGTACGACTAGCAAGACTTGAATAATCTTCAATTTGAAAT 63

DB 71511 ATTTGCTGCGTATGTGAAGAGGAGAGAGCAGAAATTCAGGATTTTCTTTACAAAT 71452

QY 64 AAACAGCTAACTCCCAATTTGCTTAAGTACGACTAGGAGGAGTTTATCATGAAGAATGTTTAAA 123

DB 71451 TTGATATTGACTTTGGAATTTGAAATAGTATAAATTTTAGAAAGTAGAGAAGTTTAAA 71392

QY 124 TGCCTTAATTTCAATAAGAGTAACCA 151

DB 71391 TTAAAAAGTTATATATTTCTTAAACACCA 71364

RESULT 34

US-11-266-748A-22663/c

Sequence 22663, Application US/11266748A

Publication No. US20060134663A1

GENERAL INFORMATION:

APPLICANT: Harkin, Paul

APPLICANT: Johnston, Patrick

APPLICANT: Mulligan, Karl

TITLE OF INVENTION: Transcriptome Microarray Technology and Methods of Using the Same

FILE REFERENCE: 55815-0102 (319189)

CURRENT APPLICATION NUMBER: US/11/266,748A

CURRENT FILING DATE: 2005-11-03

PRIOR APPLICATION NUMBER: EP 04105479.2

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105482.6

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105483.4

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105507.0

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105485.9

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105484.2

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: US 60/662,276

PRIOR FILING DATE: 2005-03-14

PRIOR APPLICATION NUMBER: US 60/700,293

PRIOR FILING DATE: 2005-07-18

NUMBER OF SEQ ID NOS: 483996

SOFTWARE: PatentIn version 3.3

SEQ ID NO 22663

LENGTH: 1071650

TYPE: DNA

ORGANISM: Homo Sapiens

US-11-266-748A-22663

Query Match 10.9%; Score 32.8; DB 8; Length 1071650;

Best Local Similarity 51.4%; Pred. No. 44;

Matches 76; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 4 ATTGACAGGCTAAATGCTTAAGTACGACTAGCAAGACTTGAATAATCTTCAATTTGAAAT 63

DB 71511 ATTTGCTGCGTATGTGAAGAGGAGAGAGCAGAAATTCAGGATTTTCTTTACAAAT 71452

QY 64 AAACAGCTAACTCCCAATTTGCTTAAGTACGACTAGGAGGAGTTTATCATGAAGAATGTTTAAA 123

DB 71451 TTGATATTGACTTTGGAATTTGAAATAGTATAAATTTTAGAAAGTAGAGAAGTTTAAA 71392

QY 124 TGCCTTAATTTCAATAAGAGTAACCA 151

DB 71391 TTAAAAAGTTATATATTTCTTAAACACCA 71364

RESULT 35

US-11-266-748A-22664/c

Sequence 22664, Application US/11266748A

Publication No. US20060134663A1

GENERAL INFORMATION:

APPLICANT: Harkin, Paul

APPLICANT: Johnston, Patrick

APPLICANT: Mulligan, Karl

TITLE OF INVENTION: Transcriptome Microarray Technology and Methods of Using the Same

FILE REFERENCE: 55815-0102 (319189)

CURRENT APPLICATION NUMBER: US/11/266,748A

CURRENT FILING DATE: 2005-11-03

PRIOR APPLICATION NUMBER: EP 04105479.2

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105482.6

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105483.4

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105507.0

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105485.9

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105484.2

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: US 60/662,276

PRIOR FILING DATE: 2005-03-14

PRIOR APPLICATION NUMBER: US 60/700,293

PRIOR FILING DATE: 2005-07-18

NUMBER OF SEQ ID NOS: 483996

SOFTWARE: PatentIn version 3.3

SEQ ID NO 22664

LENGTH: 1071650

TYPE: DNA

ORGANISM: Homo Sapiens

US-11-266-748A-22664

Query Match 10.9%; Score 32.8; DB 8; Length 247877;

Best Local Similarity 51.4%; Pred. No. 53;

Matches 76; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 4 ATTGACAGGCTAAATGCTTAAGTACGACTAGCAAGACTTGAATAATCTTCAATTTGAAAT 63

DB 52676 ATTTGCTGCGTATGTGAAGAGGAGAGAGCAGAAATTCAGGATTTTCTTTACAAAT 52617

QY 64 AAACAGCTAACTCCCAATTTGCTTAAGTACGACTAGGAGGAGTTTATCATGAAGAATGTTTAAA 123

DB 52616 TTGATATTGACTTTGGAATTTGAAATAGTATAAATTTTAGAAAGTAGAGAAGTTTAAA 52557

QY 124 TGCCTTAATTTCAATAAGAGTAACCA 151

DB 52556 TTAAAAAGTTATATATTTCTTAAACACCA 52529

```

US-11-266-748A-22664
Query Match 10.9%; Score 32.8; DB 8; Length 1071650;
Best Local Similarity 51.4%; Pred. No. 73;
Matches 76; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 4 ATTGACAGGCTAATGTAAGTACTAGTACGATGAGACTTGAATATATTCATTGTAAT 63
DDB 805354 ATTGTGCGTATGTGAAGAAGGAAGGAGGAGGAGGATTTTTCATTGTAAT 805295
QY 64 AAACAGCTAATCCCAATTTGTAAGTACTAGGAGGAGTATCATGAAGAAATGTTAAA 123
DDB 805294 TTGATATTGATTTGGAAATTTGTAAGTACTAATATTTTGAAGAGTGAAGATTTAAA 805235
QY 124 TGCTTAATTTTCAAAATGAAGTAACCA 151
DDB 805234 TTAAGAAATGATATATTTTCAACACCA 805207

RESULT 36
US-10-517-441-424/c
; Sequence 424, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOKKENS, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almuth
; APPLICANT: HOEFLE, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; PRIOR FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 424
; LENGTH: 4857
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-424

Query Match 10.9%; Score 32.6; DB 6; Length 4857;
Best Local Similarity 48.2%; Pred. No. 24;
Matches 92; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 26 GACTACGATGAGGAGTGAATATTTCTTCAATTTGAATATAACAGCTAATCCCAATTTGT 85
DDB 1075 GAAAAAATTAATAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1016
QY 86 ACCTAATAGGGAGTTTATCATGAAGAAATGTTTAAATGCTTAAATTTTCAATTAAGAG 145
DDB 1015 ACTAAAAAGGAACTAAACCTCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTA 956
QY 146 TAACGAGTGTTCACATGCTCTTAAATTAACCGCAAACTTCAATCATATAGCTGT 205

```

```

DDB 955 AAAAATACGAACCTAACTAAACCTCAAAARACCAAAACGAAATATTAGCACTCGAA 896
QY 206 AGTAGAGTGCA 216
DDB 895 AAAAAAATACA 885

RESULT 37
US-11-266-748A-224134/c
; Sequence 224134, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 224134
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-224134

Query Match 10.8%; Score 32.4; DB 8; Length 1000;
Best Local Similarity 48.9%; Pred. No. 19;
Matches 87; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 59 GAAATTAACAGCTAATCCCAATTTGTAACCTAAGAGTGTTCGCAACATGCTGTTAAATAAC 118
DDB 902 GAAACAAATATCAGTAATCTCTTTGTTCTAAACAAATTCATAATTTATTTATACATTT 843
QY 119 TTAATGCTTAAATTTTCAATAAGAGTGTTCGCAACATGCTGTTAAATAAC 178
DDB 842 TAAATATATATTTGTTTCAATGTTTGTAGTGGGCAATAATCATAAGAGATACAC 783
QY 179 CCACAACTTCAATCACTATAGTGTAGTAGTGCATTTCTGCAAGGATCCCAAGT 236
DDB 782 CAGTCGAATTAGACTTTTAAAGAAATTTACTCTCAAGACATTTGCAACAACTCCCAAGT 725

RESULT 38
US-11-266-748A-292677/c
; Sequence 292677, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A

```

```

; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 404470
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-404470

Query Match          10.8%; Score 32.4; DB 8; Length 1000;
Best Local Similarity 48.9%; Pred. No. 19;
Matches 87; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy      59  GAAATAAAACAGCTTAACTCCCAAAATTGTACCTACTAGGGGAGTTTATCATGAAGAAATGT 118
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      902  GAAACAATAATCAGTAATCCTCTTGTCTTAAACAAAATTCATAATTTATTATACATTT 843

Qy      119  TTTAAATGCTTAATTTTCAAATAGAAGTAAACAGAGTGTTGCAACATGCTGTTTAAATAAC 178
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      842  TAAAAATATTATATTTGTTTCAAATTTGTTTGTAGTGGGCAATAAAATTCATTAAGAGATACAAC 783

Qy      179  CCACACAAACTTTCACATCACTAGTCTGAGTAGAGAGTGCATTTCTGCAAGGATCCCCAGATT 236

```

Db 782 CAGTCGAATAGACTTTTAAAGAAATTACTCTCAAGACATTTGCAACAATCCAAAGT 725

RESULT 41

US-11-266-748A-475516
; Sequence 475516, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 475516
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-475516

Query Match 10.8%; Score 32.4; DB 8; Length 1000;
Best Local Similarity 48.9%; Pred. No. 19;
Matches 87; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 59 GAAATAACAGCTAACTCCCAAATTTGACCTTAAGTGGGAGTTTATCATGAAGAAATGT 118
Db 99 GAAACAAATATCAGTAATCTCTTTGTTCTTAACAAATAATTCATAATTTATACATTT 158
QY 119 TTAATGCTTTAAATTTTCAATAAGAGTAACAGAGTGTTCGAACATGCTGTTTAAATAAC 178
Db 159 TAAATATTATTTGTTTCAAAATGTTGTAGTGGGCAATAAATCATAAAGAGATACAAC 218
QY 179 CCGCAAACTTCAATCACTATAGCTGTAGTAGAGTGCATTTGCAAGATCCCAAGT 236
Db 219 CAGTCGAATAGACTTTTAAAGAAATTACTCTCAAGACATTTGCAACAATCCAAAGT 276

RESULT 42

US-11-266-748A-71950/c
; Sequence 71950, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 71950
; LENGTH: 1039
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-71950
Query Match 10.8%; Score 32.4; DB 8; Length 1039;
Best Local Similarity 48.9%; Pred. No. 19;
Matches 87; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 59 GAAATAACAGCTAACTCCCAAATTTGACCTTAAGTGGGAGTTTATCATGAAGAAATGT 118
Db 957 GAAACAAATATCAGTAATCTCTTTGTTCTTAACAAATAATTCATAATTTATACATTT 898
QY 119 TTAATGCTTTAAATTTTCAATAAGAGTAACAGAGTGTTCGAACATGCTGTTTAAATAAC 178
Db 897 TAAATATTATTTGTTTCAAAATGTTGTAGTGGGCAATAAATCATAAAGAGATACAAC 838
QY 179 CCGCAAACTTCAATCACTATAGCTGTAGTAGAGTGCATTTGCAAGATCCCAAGT 236
Db 837 CAGTCGAATAGACTTTTAAAGAAATTACTCTCAAGACATTTGCAACAATCCCAAGT 780

RESULT 43

US-11-266-748A-124761
; Sequence 124761, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; APPLICANT: Johnston, Patrick
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 124761
; LENGTH: 1039
; TYPE: DNA

```
; ORGANISM: Homo Sapiens
US-11-266-748A-124761

Query Match      10.8%; Score 32.4; DB 8; Length 1039;
Best Local Similarity 48.9%; Pred. No. 19;
Matches 87; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 59 GAAATAAACAGCTAACTCCCAAAATGTACCTAACTAGGGAGTTTATCATGAAGAAATGT 118
DB 83 GAAACAAATATACGTAATCCTCTTTGTTCTAAACAAAATTCATATTTATATACATTT 142
QY 119 TTAATGCTTAATTTTCAAAATGAAGTAACACAGAGTGTGCAACATGCTGTTAAATAAC 178
DB 143 TAAATATATATGTTTCAATGTTGTAGTGGGCAATTAATCATAAAGAGATACAAAC 202
QY 179 CCGCAAACTTCAATCACTATAGCTGTAGTAGAGTGCAATTCGCAAGGATCCCAAGT 236
DB 203 CAGTCGAATTAGACTTTTAAAGRATTACTCTCAAAGACATTCGCAACAACTCCAAAGT 260

RESULT 44
US-11-266-748A-24100/c
; Sequence 24100, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 24100
; LENGTH: 4862
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-24100

Query Match      10.8%; Score 32.4; DB 8; Length 4862;
Best Local Similarity 48.9%; Pred. No. 27;
Matches 87; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 59 GAAATAAACAGCTAACTCCCAAAATGTACCTAACTAGGGAGTTTATCATGAAGAAATGT 118
DB 4781 GAAACAAATATACGTAATCCTCTTTGTTCTAAACAAAATTCATATTTATATACATTT 4722
QY 119 TTAATGCTTAATTTTCAAAATGAAGTAACACAGAGTGTGCAACATGCTGTTAAATAAC 178
DB 4721 TAAATATATATGTTTCAATGTTGTAGTGGGCAATTAATCATAAAGAGATACAAAC 4662
QY 179 CCGCAAACTTCAATCACTATAGCTGTAGTAGAGTGCAATTCGCAAGGATCCCAAGT 236
DB 4661 CAGTCGAATTAGACTTTTAAAGRATTACTCTCAAAGACATTCGCAACAACTCCAAAGT 4604
```

```
RESULT 45
US-11-217-529-166179/c
; Sequence 166179, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 166179
; LENGTH: 8391
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-166179

Query Match      10.8%; Score 32.4; DB 8; Length 8391;
Best Local Similarity 58.2%; Pred. No. 31;
Matches 57; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 40 CTTGAAATATTTCTTCATTGAAATAAACAGCTAACTCCCAAAATGTACTTAAGGGA 99
DB 5463 CTTTATATATTTCTTAATAAAATATAAAGAAATTAATAATAATAATAATAATAA 5404
QY 100 GTTTATCATGAGAAATGTTTAAATGCTTAAATTTTCA 137
DB 5403 CAATATAATAATAATAATAATAATAATAATAATAACAA 5366

RESULT 46
US-11-266-748A-22572
; Sequence 22572, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 22572
; LENGTH: 31857
; TYPE: DNA
```



```

; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 135265
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-135265

Query Match      10.7%; Score 32.2; DB 8; Length 1218;
Best Local Similarity 59.1%; Pred. No. 22;
Matches 55; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 42 TGAATAATTCATTTGAAATAAACAGCTAACTCCCAAAATGTACCTAACTAGGGGAGT 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 876 TGAGGTATATTTCTTTTAAACATTAGATACTAGCAAAATATATATACCAGGAGT 935
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 102 TTATCATGAAGAAATGTTTAAATGCTTAATTTT 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 936 TTCAAATTAATATATTACAGGGCTTGCTTTT 968
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 50
US-11-266-748A-30890/c
; Sequence 30890, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266.748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 30890
; LENGTH: 4388
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-30890

Query Match      10.7%; Score 32.2; DB 8; Length 4388;
Best Local Similarity 59.1%; Pred. No. 30;
Matches 55; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 42 TGAATAATTCATTTGAAATAAACAGCTAACTCCCAAAATGTACCTAACTAGGGGAGT 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3210 TGAGGTATATTTCTTTTAAACATTAGATACTAGCAAAATATATATACCAGGAGT 3151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 102 TTATCATGAAGAAATGTTTAAATGCTTAATTTT 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3150 TTCAAATTAATATATTACAGGGCTTGCTTTT 3118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Search completed: July 18, 2006, 00:52:40
Job time : 128 secs

```

This Page Blank (uspio)